

An extensive resonant normal form for an arbitrary large Klein-Gordon model

Simone Paleari and Tiziano Penati

October 9, 2014

Abstract

We consider a finite but arbitrarily large Klein-Gordon chain, with periodic boundary conditions. In the limit of small couplings in the nearest neighbor interaction, and small (total or specific) energy, a high order resonant normal form is constructed with estimates uniform in the number of degrees of freedom. In particular, the first order normal form is a generalized discrete nonlinear Schrödinger model, characterized by all-to-all sites coupling with exponentially decaying strength.

Keywords: Extensivity, resonant normal form, Klein-Gordon model, anticontinuum limit, Thermodynamic limit, generalized dNLS model.

1 Introduction and statement of the results

In the present paper, along the lines of [3, 4, 13, 14, 18], we keep on investigating the development and application of a perturbation theory for Hamiltonian systems with an arbitrarily large number of degrees of freedom, and in particular in the thermodynamic limit. Indeed, motivated also by the problems arising in the foundations of Statistical Mechanics, we want to consider large systems (e.g. for a model of a crystal the number of particles should be of the order of the Avogadro number) with non vanishing energy per particle (which corresponds to a non zero temperature in the physical model).

Since we are interested in the low temperature regime (aiming for example at some rigorous results of the classical mechanics description of the behavior of the specific heats in such a regime), it is foreseeable the use of perturbation theory to exploit the presence of a small parameter like the specific energy. Unfortunately, it is a well known limit of the classical results of this theory (like KAM or Nekhoroshev theorem) to suffer a bad dependence on the number of degrees of freedom, often resulting in void or non applicable statements in the thermodynamic limit.

In the recent papers [4, 14, 18] it has been possible to prove, for the first time, the existence of an approximate conserved quantity, independent of the Hamiltonian, exactly in the thermodynamic limit, thus with uniform estimates in the number of degrees of freedom and non vanishing specific energy.

In the present work we make progress in the above mentioned research program, with a result in the direction of a normal form construction, rather than in that of approximate conserved quantities. Such a construction is shown to hold both in a regime of small total energy and in a regime of small specific energy; moreover we are able to completely control

the dependence on the number of degrees of freedom, thanks to the ideas and techniques used in [14], that we here extend also to cover the normal form algorithm.

We consider a Klein-Gordon model as described by the following Hamiltonian

$$H(x, y) = \frac{1}{2} \sum_{j=1}^N [y_j^2 + x_j^2 + a(x_{j+1} - x_j)^2] + \frac{1}{4} \sum_{j=1}^N x_j^4, \quad x_0 = x_N, \quad y_0 = y_N, \quad (1)$$

i.e. a finite chain of N degrees of freedom and periodic boundary conditions.

Our main result holds in the limit of small coupling (the constant a in the Hamiltonian), and small (both total and specific) energy. A simplified statement of our normal form construction could be the following (see Proposition 3.1 and Theorem 3.1 for the complete ones)

Theorem *There exist C_1 and C_2 , such that for every N , every small enough value of the coupling constant a , every integer $r < C_1/a$, and (total or specific) energy less than C_2/r^2 , there exists an analytic canonical transformation, under which the Hamiltonian (1) takes the form*

$$H^{(r)} = H_\Omega + Z_0 + \dots + Z_r + P^{(r+1)}, \quad \{H_\Omega, Z_s\} = 0 \quad \forall s \in \{0, \dots, r\}.$$

with H_Ω a system of N identical harmonic oscillators, Z_s homogeneous polynomials of order $2s + 2$, $P^{(r+1)}$ a remainder of order $2r + 4$ and higher.

The first aspect we need to remark here is that, in order to control as in [14] the dependence on N in the whole perturbation construction, we exploit the fact that the Hamiltonian is extensive, i.e. the energy of the system is, roughly speaking, proportional to N . The extensivity is the result of two general properties of our (and similar) model, which are introduced and discussed in general in Section 2. The first one is the translational invariance, that we formalize through what we call cyclic symmetry (see definition 2.1 and in general subsection 2.1): as it happens in a cyclic group, we exploit this discrete symmetry by introducing the idea of a generator f (a “seed” in our terminology, see (3)) of a translation-invariant function F . The second property is the short interaction range (see subsection 2.2): actually (1) in the original variables possesses a finite range (nearest neighbor) interaction, which is immediately replaced, in the perturbation construction, by an infinite range interaction with an exponentially decaying strength with the distance. These properties of cyclic symmetry and exponential decay in the interaction range are explicitly controlled in particular for the whole transformed Hamiltonian in Theorem 3.1, allowing us to obtain estimates uniform with N . Indeed, it is possible to implement the whole perturbation scheme at the level of the seeds, whose norms (due to the short range interaction) are independent¹ of N . In our opinion this represents the main original aspect in the use of the translational discrete symmetry inside a perturbation construction for an Hamiltonian chain. The possibility to preserve a discrete symmetry while performing a normal form construction is surely not completely new: as an

¹Although it is not our aim to give here a formal definition of “extensivity”, we could associate the concept of being “extensive” for a function F , with the following two properties: for F to be cyclically symmetric and for its generator/seed f to have a norm independent of N . Indeed this could be a rather general characterization, which is meaningful for every function. And if in particular F represents the potential of an interaction force, then, since the independence of $\|f\|$ with respect to N is equivalent to the sufficiently fast decay of the interaction with the distance among sites, one could recover the usual idea of extensivity given by translation invariance plus short range interaction.

example, we could mention [24, 25], where the dyhedral group symmetries have been successfully exploited to show the Liouville integrability of the Birkhoff-Gustavson normal form for a periodic FPU lattice. However, the benefits of our strategy of working at the level of the “seeds” in the perturbation scheme go beyond the information on the structure of the normal form: it is the key ingredient to even provide estimates, also with a sharp dependence on the number of degree of freedom of the system.

Another aspect to remark is the validity of the normal form Theorem both in specific and total energy regime: and if in particular we restrict to the small total energy regime some dynamical information can be immediately obtained, since it is possible to deduce a somewhat complementary result to that of [14]. Indeed, in Corollary 3.1 (see Section 3), we have provided a long time adiabatic invariance of the ℓ^2 norm H_Ω in the classical sense (i.e. not in the probabilistic formulation of [14]). An analogue result in the form of a dynamical application of our Theorem in the specific energy regime is instead a much more difficult task, and reasonable results in that direction are still missing.

It is nevertheless worth to put more into evidence other potential advantages of the present normal form construction, which gives more information on the structure of the Hamiltonian. Indeed, in Section 4 we concentrate on the low order normal form $H_\Omega + Z_0 + Z_1$ (hence choosing $r = 1$ in the Theorem), which represents a generalized discrete nonlinear Schrödinger (GdNLS) chain, characterized by all-to-all sites couplings, both in the linear and nonlinear terms, with exponential decay of the coefficients with the distance between sites. Since it turns out that such a decay is given by powers of the small coupling constant a , a truncation of this normal form results in the usual dNLS model, which is well known to provide a leading order approximation of the KG evolution, both in the continuum limit (as for example in [1]) and in the anticontinuum limit (as discussed for example in [20] or formally used in the modulational instability description like in [9, 10]).

We think that such a GdNLS model can be interesting by itself: indeed, due to the approach used in its construction, it contains the right normal form for several different regimes with respect to the relative smallness between the two small parameters, the coupling and the energy.

We also observe that, in the regime of small energy, the result is valid also in the case of the soft nonlinearity, i.e. with a minus in front of the quartic term; we stress that, in such a case, the second order normal form $H_\Omega + Z_0 + Z_1 + Z_2$ could be seen as a perturbation (again due to the all-to-all sites couplings) of the cubic-quintic dNLS (see, e.g., [5, 7]), with competing nonlinearities.

We close the comments on the GdNLS with the remark that it could be the starting point for several applications which concern the Klein-Gordon model, like the variational approximations for breather solutions (see, e.g., [7, 8]), the approximation of the small amplitude Cauchy problem, the existence and linear stability of multibreathers (see, e.g., [15, 16, 20, 21]). Moreover there are several recent works on models with more than first neighbor interactions or with different nonlinearities, like [6, 17, 23, 26] where spatially localized periodic orbits, as breathers or multibreathers, are studied: with respect to this, we expect that the approach proposed in the present paper can be suitably extended in these more general models, leading to different GdNLS-like normal forms. As an example, in a subsequent paper [19], following also some ideas from [2] and still in the spirit of dealing with the case $N < \infty$ (see [22] for a dNLS study), we apply this normal form construction to a mixed KG/FPU model in order to provide a result of long time approximation of the dynamics close to site-symmetric breather solutions of a GdNLS model, whose existence and stability are proved as intermediate step.

As a last remark on the potential benefit of the present normal form and of the techniques developed and used, we recall the realm of numerical and/or symbolic computations. Indeed it would be interesting to figure out how to exploit our formalization of the extensivity in infinite systems (e.g. some classes of PDEs) in order to perform algebraic manipulations for these type of systems.

The paper is structured as follows. In Section 2 we recall the formalization of the physical properties of the model, including some results of [14] and adding the control of Hamiltonian vector fields. In Section 3 we present and further comment the results of the paper in a more detailed and complete form. In Section 4 we discuss the GdNLS normal form. In Section 5 the formal construction is presented and most of the proofs of the estimates are given, with further technical proof deferred to the Appendix 6.

2 A formalization of extensivity: cyclic symmetry and short range interaction

This Section is devoted to the formalization of two fundamental, though quite general, properties of the KG chain, i.e. the discrete translation invariance that we call *cyclic symmetry*, and the short interaction range. Most of the definitions and the results have been already introduced in [13,14]: we repeat them here in Subsections 2.1 and 2.2, with a hopefully more terse exposition, to let the present paper be self-contained. In Subsection 2.3 we add some new results, i.e. the treatment of vector fields, which were not necessary in our previous papers.

The system under investigation, i.e. the finite but large KG chain described by the Hamiltonian (1), possesses some general properties shared by a lot of many particle systems, as discussed in [14]: they are characterized by two-body conservative forces with smooth potential which are invariant with respect to rotations and/or translations. These properties are quite general ones.

Since we will apply our construction on the KG chain, we restrict our attention to a system of identical particles on a d -dimensional lattice, with a short or even finite range interaction². Moreover it is enough to know the local interaction of a particle with its neighbors and the complete Hamiltonian is the sum of the contribution of every particle to both the kinetic and the potential energy. We thus have the presence of both such a cyclic symmetry, and of a short range interaction potential. These properties characterize and somewhat formalize the fact that the Hamiltonian is *extensive*, i.e. it is proportional to the number of degrees of freedom N . Functions possessing the same extensivity property of the Hamiltonian are particularly relevant.

2.1 Cyclic symmetry

We consider the simplified model of a finite one dimensional lattice with periodic boundary conditions and finite range interactions (nearest neighbors). We denote by x_j, y_j the position and the momentum of a particle, with $x_{j+N} = x_j$ and $y_{j+N} = y_j$ for any j .

²We recall that by “finite range interaction” we mean an interaction which, for any particle, involves only a finite number of neighbors, independent of N ; a “short range interaction” instead may involve even an all-to-all interaction provided its strength decays fast enough with the distance.

Cyclic symmetry. We formalize one of the ingredient of the extensivity, i.e. discrete translation invariance, by using the idea of *cyclic symmetry*. In [13, 14] we decided to use such a new and nonstandard terminology both to remind that the associated group is a cyclic one, and also because we think we exploited the invariance in a novel way. The *cyclic permutation* operator τ , acting separately on the variables x and y , is defined as

$$\tau(x_1, \dots, x_N) = (x_2, \dots, x_N, x_1) , \quad \tau(y_1, \dots, y_N) = (y_2, \dots, y_N, y_1) . \quad (2)$$

We extend its action on the space of functions as

$$(\tau f)(x, y) = f(\tau(x, y)) = f(\tau x, \tau y) .$$

Definition 2.1 *We say that a function F is cyclically symmetric if $\tau F = F$.*

In order to further exploit the symmetry, we now try to formalize the idea that, from the point of view of information content, there is a lot of redundancy in a cyclically symmetric function, i.e. it can be reconstructed from a “smaller” object. We thus introduce an operator, indicated by an upper index \oplus , acting on functions: given a function f , a new function $F = f^\oplus$ is constructed as

$$F = f^\oplus := \sum_{l=1}^N \tau^l f . \quad (3)$$

We shall say that $f^\oplus(x, y)$ is generated by the *seed* $f(x, y)$. We will try to use the convention of denoting cyclically symmetric functions with capital letters and their seeds with the corresponding lower case letter.

Lemma 2.1 (see [14]) *The following holds:*

1. *given a seed f , then for $F = f^\oplus$ one has $\tau F = F$;*
2. *given a function F , such that $\tau F = F$, then there exist a (not unique) seed f such that $F = f^\oplus$;*
3. *for any integers s_1, s_2 , $(f_1 + f_2)^\oplus = (\tau^{s_1} f_1 + \tau^{s_2} f_2)^\oplus$;*
4. *the Poisson bracket between two cyclically symmetric functions is also cyclically symmetric, i.e. we may write $h^\oplus = \{f^\oplus, g^\oplus\}$. A candidate seed is $h = \{f, g^\oplus\}$.*

It is worth to stress that property 4 of the above Lemma is the one which allows to perform the normal form construction by preserving the cyclic symmetry. From a purely formal point of view, the compatibility of the discrete translation invariance with a canonical perturbation construction is not new, since the possibility to perform the Lie-transform normalization by preserving a discrete (and symplectic) symmetry is a well known fact (see for example [24, 25] and references therein). Once again we bring this fact down the level of seeds to exploit it further.

Polynomial norms. Let $f(x, y) = \sum_{j,k} f_{j,k} x^j y^k$ be a homogeneous polynomial of degree s in x, y . Given a positive R , we define its polynomial norm as

$$\|f\|_R := R^s \sum_{|j|+|k|=s} |f_{j,k}|. \quad (4)$$

If R represents the radius of the ball centered in the origin of the phase space, endowed for example with the euclidean norm, then one would have

$$|f(x, y)| \leq \sup_{\|(x,y)\| \leq R} |f(x, y)| \leq \|f\|_R.$$

Norm of a cyclically symmetric function. Assume now that we are equipped with a norm for our functions $\|\cdot\|$, e.g. the above defined polynomial norm. We introduce a corresponding norm $\|\cdot\|^\oplus$ for a cyclically symmetric function $F = f^\oplus$ by defining

$$\|F\|^\oplus = \|f\|, \quad (5)$$

i.e. we actually measure the norm of the seed. An obvious remark is that the norm so defined depends on the choice of the seed, but this will be harmless in the following. The relevant facts are the following:

Lemma 2.2 (see [14]) *It holds:*

1. for any s one has $\|\tau^s f\| = \|f\|$;
2. the inequality $\|f^\oplus\| \leq N \|f^\oplus\|^\oplus$ holds true for any choice of the seed.

This is particularly useful when the norm of the seed turns out to be independent of N , since in such a case the dependence on the number of degrees of freedom is completely factorized and totally under control. Moreover one could verify if a function is “extensive” by checking if it is cyclically symmetric and with its cyclic norm independent of N . We remark that, if the function under consideration is an interaction potential, then obviously this second property is equivalent to the interaction range³ being short: indeed, it is possible to get a seed whose norm (4) is independent on N also in the case of an all-to-all interaction like

$$\sum_{i,j} a_{i,j} (x_i - x_j)^2,$$

provided some suitable decay of the coefficients $|a_{i,j}|$ with the distance $|i - j|$.

Circulant matrices. When we deal with particular functions which are quadratic forms, the cyclic symmetry assumes a particular form. Let us thus restrict our attention to the harmonic part of the Hamiltonian: it is a quadratic form represented by a matrix A

$$H_0(x, y) = \frac{1}{2} y \cdot y + \frac{1}{2} A x \cdot x. \quad (6)$$

³more on this point in Subsection 2.2

If the Hamiltonian H_0 is cyclically symmetric, then $H_0 = h_0^\oplus$. This implies that A commutes with the matrix τ representing the cyclic permutation (2)

$$\tau_{ij} = \begin{cases} 1 & \text{if } i = j + 1 \pmod{N}, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

We remark that the matrix τ is orthogonal and generates a cyclic group of order N with respect to the matrix product.

We recall the following

Definition 2.2 *A matrix $A \in \text{Mat}_{\mathbb{R}}(N, N)$ is said to be circulant if*

$$A_{j,k} = a_{(k-j) \pmod{N}}. \quad (8)$$

Actually, the set of circulant matrices is a subset of Toeplitz matrices, i.e. those which are constant on each diagonal. For a comprehensive treatment of circulant matrices, see, e.g., [11]. We just remind some properties that will be useful later.

1. The set of $N \times N$ circulant matrices is a real vector space of dimension N , and a basis is given by the cyclic group generated by τ .
2. The set of matrices which commute with τ coincides with the set of circulant matrices.
3. The set of eigenvalues of a circulant matrix is the Discrete Fourier Transform of the first row of the matrix and viceversa.
4. Let $M^2 = A$, where A is circulant; then M is circulant, too. Moreover, from the definition of $M := \sqrt{A}$, it follows that if A is symmetric, then M is also symmetric.

In our problem the cyclic symmetry of the Hamiltonian implies that the matrix A of the quadratic form is circulant. Obviously it is also symmetric, so that the space of matrices of interest to us has dimension $\lfloor \frac{N}{2} \rfloor + 1$. Indeed, a circulant and symmetric matrix is completely determined by $\lfloor \frac{N}{2} \rfloor + 1$ elements of its first line.

2.2 Interaction range

Besides the translation invariance, usually the second ingredient for the formalization of extensivity is the sufficiently fast decay of the interaction strenght⁴, which is equivalent to the independence on N of the cyclic norm of the interaction potential. We give here some definitions and properties at the level of the functions' seeds. We restrict our analysis to the set of polynomial functions. We start with some definitions. Let us label the variables as x_l, y_l with $l \in \mathbb{Z}$, and consider a monomial $x^j y^k$ (in multiindex notation).

Definition 2.3 *We define the support $S(x^j y^k)$ of the monomial and the interaction distance $\ell(x^j y^k)$ as follows: considering the exponents (j, k) we set*

$$S(x^j y^k) = \{l : j_l \neq 0 \text{ or } k_l \neq 0\}, \quad \ell(x^j y^k) = \text{diam}(S(x^j y^k)). \quad (9)$$

We say that the monomial is left aligned in case $S(x^j y^k) \subset \{0, \dots, \ell(x^j y^k) - 1\}$.

⁴At least for those function for which the concept makes sense, i.e. those giving an interaction potential.

The definitions above is extended to a homogeneous polynomial f by saying that $S(f)$ is the union of the supports of all the monomials in f , and that f is left aligned if all its monomials are left aligned. The relevant property is that if \tilde{f} is a seed of a cyclically symmetric function F , then there exists also a left aligned seed f of the same function F : just left align all monomials in \tilde{f} .

Short range (exponential decay of) interaction. Although it is not necessary for the interaction to be short, we consider the case of an exponential decay of the interaction strength, since in our case this is the property which holds. For the seed f of a function consider the decomposition

$$f(z) = \sum_{m \geq 0} f^{(m)}(z) , \quad f^{(m)}(z) = \sum_{\ell(k) \leq m} f_k z^k , \quad (10)$$

assuming that every $f^{(m)}$ is left aligned.

Definition 2.4 *The seed f (of a cyclically symmetric function) is of class $\mathcal{D}(C_f, \sigma)$ if*

$$\|f^{(m)}\|_1 \leq C_f e^{-\sigma m} , \quad C_f > 0 , \sigma > 0 . \quad (11)$$

Remark 2.1 *It is immediate to notice that when C_f does not depend on N , then*

$$\|f\|_1 \leq \sum_{m \geq 0} \|f^{(m)}\|_1 \leq C_f \sum_{m \geq 0} e^{-\sigma m} = \frac{C_f}{1 - e^{-\sigma}} ,$$

hence $\|f\|_1$ does not grow with N .

2.3 Hamiltonian vector fields

We introduce here some definitions and some results concerning Hamiltonian vector fields, their Lie derivatives, and the control of their norms. This part is completely absent in [14] since in such a paper all the perturbation construction is performed at the level of the Hamiltonian functions and not at the level of the vector fields.

We consider, as an Hamiltonian, a cyclically symmetric function F with seed f ; we will make use of the common notation⁵ $X_F = (X_1, \dots, X_N, X_{N+1}, \dots, X_{2N})$ to indicate the associated Hamiltonian vector field $J\nabla F$, with J given by the Poisson structure. The first easy, but important, result is that also the Hamiltonian vector field inherits, in a particular form, the cyclic symmetry; a possible choice for the equivalent of the seed turn out to be the pair (X_1, X_{N+1}) , i.e. the first and the $(N+1)^{\text{th}}$ components of the vector. This fact, which will be more clear thanks to the forthcoming Lemma 2.3, allows us to define in a reasonable and consistent way the following norm

$$\|X_F\|_R^\oplus := \|X_1\|_R + \|X_{N+1}\|_R . \quad (12)$$

⁵For an easier notation we drop the Hamiltonian F in the indexes of the components of the vector field.

Lemma 2.3 *Given $F = f^\oplus$, for the components of its Hamiltonian vector field X_F we have⁶*

$$\begin{aligned} X_j &= \tau^{j-1} X_1 \\ X_{N+j} &= \tau^{j-1} X_{N+1} \end{aligned} \quad j = 1, \dots, N. \quad (13)$$

Moreover, it holds

$$\|X_F\|_R^\oplus = \sum_{l=1}^{2N} \left\| \frac{\partial f}{\partial z_l} \right\|_R. \quad (14)$$

proof: We start by observing the following identity about the commutation properties of partial derivative and cyclic permutation defined in (2):

$$\frac{\partial}{\partial x_j} [f \circ \tau^l] = \tau^{-l} \left[\frac{\partial f}{\partial x_{j+l}} \right],$$

where, as usual, all the index for the variables are meant modulo N , independently for each set x and y . The similar relation holds for the partial derivatives with respect to the y variables.

Using that $F = \sum_l \tau^l f$, and using the above relation to “extract” a permutation τ^{1-j} , we have for $j = 1, \dots, N$

$$\begin{aligned} X_j &\equiv \frac{\partial F}{\partial y_j} = \sum_l \frac{\partial}{\partial y_j} [\tau^l f] = \sum_l \tau^{j-1} \left[\frac{\partial}{\partial y_1} (\tau^{l+j-1} f) \right] = \\ &= \tau^{j-1} \sum_m \frac{\partial}{\partial y_1} [\tau^m f] = \tau^{j-1} \frac{\partial F}{\partial y_1} \end{aligned}$$

which gives the first of (13). Analogously for X_{N+j} . Concerning the equality (14) one uses again the commutation properties stated at the beginning of the proof, and then the invariance of the polynomial norm $\|\cdot\|_R$ under the action of τ . \square

Definition 2.5 *We denote with \mathcal{P} the phase $(\mathbb{R}^{2N}, \|\cdot\|)$, endowed by either the euclidean norm (ℓ^2) or the supremum norm (ℓ^∞). When necessary, we will specify the norm used with a subscript, i.e. \mathcal{P}_2 with $\|\cdot\|_2$ and \mathcal{P}_∞ with $\|\cdot\|_\infty$.*

It is easy to check that, when dealing with “local” potentials like $V(x) = \sum_j \frac{1}{2r} x_j^{2r}$, the corresponding Hamiltonian field X_V fulfills

$$\|X_V(x, y)\| \leq \|x\|^{2r-1};$$

with both the above introduced norms. Our aim is to generalize the above estimate to cyclically symmetric Hamiltonian fields X_F with $\|X_F\|_1^\oplus < \infty$. To motivate the forthcoming definition (16), we remark that for any polynomial vector field $X(z)$ of degree r there exists a r -linear operator $\tilde{X}(z_1, \dots, z_r)$ such that

$$X(z) = \tilde{X}(z, \dots, z). \quad (15)$$

⁶An immediate consequence of (13) is that, defining the norm of the vector field as the sum of its components (i.e. a finite ℓ^1 norm), we would get $\|X_F\|_R = N \|X_F\|_R^\oplus$, which in turn justify the definition (12), and make it consistent with our previous definition (5).

Definition 2.6 For a polynomial vector field X of degree r define an “operator norm”

$$\|X\|_{\text{op}} := \sup_{\|z\| \neq 0} \frac{\|X(z)\|}{\|z\|^r}, \quad (16)$$

where, on the right hand side, all the $\|\cdot\|$ can be either $\|\cdot\|_2$ or $\|\cdot\|_\infty$.

The following result, whose proof is deferred to the Appendix 6.2, gives the above claimed control of the cyclically symmetric Hamiltonian fields. We stress that it is valid both in \mathcal{P}_2 and in \mathcal{P}_∞ .

Proposition 2.1 Let f be an homogeneous polynomial of degree $r+1$ with $r \geq 1$ and $F = f^\oplus$ the cyclically symmetric Hamiltonian generated by f . Then it holds true

$$\|X_F\|_{\text{op}} \leq \|X_F\|_1^\oplus. \quad (17)$$

We close this Section with a statement (whose proof is also in the Appendix, see 6.3) providing the estimate on the Hamiltonian vector field of a function of class $\mathcal{D}(C_f, \sigma)$.

Lemma 2.4 Let F be cyclically symmetric homogeneous polynomials of degree r and let its seed f be of class $\mathcal{D}(C_f, \sigma)$; then

$$\|X_F\|_R^\oplus \leq 4rR^{r-1} \frac{C_f}{(1 - e^{-\sigma})^2}. \quad (18)$$

3 Results

In this section we present the extensive resonant normal form Theorem for the Hamiltonian (1); in the subsequent Section 4 we will add some preliminary applications⁷ of such a result, exhibiting a generalized dNLS as a first order normal form of (1).

In order to present the result we split the Hamiltonian (1) as a sum of its quadratic and quartic parts $H = H_0 + H_1$, where

$$H_0(x, y) := \frac{1}{2} \sum_{j=1}^N [y_j^2 + x_j^2 + a(x_j - x_{j-1})^2], \quad H_1(x, y) := \frac{1}{4} \sum_{j=1}^N x_j^4. \quad (19)$$

3.1 Normalization of the quadratic part

The first step is the application of the same initial linear transformation used in [13, 14] to give the quadratic part a resonant normal form. This is a preliminary operation which is absolutely necessary in order to “prepare” the Hamiltonian H for the forthcoming perturbation algorithm. As widely discussed in the above cited papers, such a normalization can be implemented using different approaches. We recall here a simplified statement of the corresponding one of [14]. Let us recall the matrix A introduced in (6)

$$A = (1 + 2a) \left[\mathbb{I} - \mu(\tau + \tau^\top) \right], \quad \text{with} \quad \mu := \frac{a}{1 + 2a}, \quad (20)$$

⁷In a forthcoming paper [19] we will exploit further the present theorem for some Breathers stability result.

which is clearly circulant and symmetric (recall τ is the permutation matrix generating (2)), and gives a finite range interaction, in the form of a μ small⁸ perturbation of the identity. We also introduce the constant frequency Ω as the average of the square roots of the eigenvalues of A (actually, the frequencies of the linearized oscillations). Let us introduce the exponent

$$\sigma_0 := -\ln(2\mu) , \quad (21)$$

and take any positive $\sigma_1 < \sigma_0$. We have

Proposition 3.1 (see [14]) *For $\mu < 1/2$, the canonical linear transformation $q = A^{1/4}x$, $p = A^{-1/4}y$ gives the Hamiltonian H_0 the particular resonant normal form*

$$H_0 = H_\Omega + Z_0 , \quad \{H_\Omega, Z_0\} = 0 \quad (22)$$

with H_Ω and Z_0 cyclically symmetric with seeds

$$h_\Omega = \frac{\Omega}{2}(q_1^2 + p_1^2) , \quad \zeta_0 \in \mathcal{D}(C_{\zeta_0}(a), \sigma_0) ,$$

and transform H_1 into a cyclically symmetric function with seed

$$h_1 \in \mathcal{D}(C_{h_1}(a), \sigma_1) .$$

We stress that it is the above linear transformation which introduces in a natural way, both in Z_0 and in H_1 , the interaction among all sites, with an exponential decay with respect to their distance. Differently from the quadratic interaction Z_0 , the seed h_1 cannot⁹ preserve the same exponential decay rate of the linear transformation; however, as claimed in the above Proposition, it is possible to show that $h_1 \in \mathcal{D}(C_{h_1}(a), \sigma_1)$ for any $\sigma_1 < \sigma_0$. We here make the choice

$$\sigma_1 := \frac{1}{2}\sigma_0 , \quad (23)$$

in order to explicitly relate σ_1 to the small natural parameter a of the model.

3.2 Normal Form Theorem

With the Hamiltonian transformed by means of the above Proposition into the form

$$H = H_\Omega + Z_0 + H_1 , \quad (24)$$

we are now ready to state the main Theorem. We only anticipate that the idea is to perform, by using the Lie transform algorithm in the form explained in [12], r normalizing steps, provided $r < r_*(\mu)$. As expected, the maximum number $r_*(\mu)$ of steps allowed increases when μ decreases. Moreover, given μ and r , the normalizing canonical transformation is well defined in a (small) neighborhood B_R of the origin, where $R < R_*(r, \mu)$. Although this canonical transformation preserves the extensive nature of the system, at any step one has to lose a bit of the exponential decay of the interactions involved in the Hamiltonian.

⁸ μ is essentially proportional to the natural small coupling a , and is always less than one half since we consider positive values of a .

⁹We mention here that this loss of the exponential decay is a consequence of the requirement that the seed h_1 has to be left aligned. Indeed it is actually possible to keep $h_1 \in \mathcal{D}(\cdot, \sigma_0)$, but with a different expansion of $h_1 = \sum_l h_1^{(l)}$: namely if the support $S(h_1^{(l)})$ is not left aligned but “symmetrically aligned” around the 0-th site (see also Section 4).

Theorem 3.1 Consider the Hamiltonian $H = h_\Omega^\oplus + \zeta_0^\oplus + h_1^\oplus$ with seeds $h_\Omega = \frac{\Omega}{2}(x_0^2 + y_0^2)$, the quadratic term ζ_0 of class $\mathcal{D}(C_{\zeta_0}, \sigma_0)$ with $\zeta_0^{(0)} = 0$, and the quartic term h_1 of class $\mathcal{D}(C_{h_1}, \sigma_1)$. Pick a positive $\sigma_* \in [\max(\ln(4), \sigma_0/4), \sigma_1)$; then there exist positive γ , μ_* and C_* such that for any positive integer r satisfying

$$r < \frac{1}{2} \left(\frac{\mu_*}{\mu} \right), \quad (25)$$

there exists a finite generating sequence $\mathcal{X} = \{\chi_1^\oplus, \dots, \chi_r^\oplus\}$ of a Lie transform such that $T_{\mathcal{X}}H^{(r)} = H$ where $H^{(r)}$ is a cyclically symmetric function of the form

$$H^{(r)} = H_\Omega + \mathcal{Z} + P^{(r+1)}, \quad \begin{aligned} \mathcal{Z} &:= Z_0 + \dots + Z_r \\ L_\Omega Z_s &= 0, \quad \forall s \in \{0, \dots, r\}, \end{aligned} \quad (26)$$

with Z_s of degree $2s + 2$ and $P^{(r+1)}$ a remainder starting with terms of degree equal or bigger than $2r + 4$.

Moreover, defining $C_r := 64r^2 C_*$ and $\sigma_j := \sigma_1 - \frac{j-1}{r}(\sigma_1 - \sigma_*)$, the following statements hold true:

(i) the seed χ_s of \mathcal{X}_s is of class $\mathcal{D}(C_r^{s-1} \frac{C_{h_1}}{\gamma_s}, \sigma_s)$.

(ii) the seed ζ_s of Z_s is of class $\mathcal{D}(C_r^{s-1} \frac{C_{h_1}}{s}, \sigma_s)$.

(iii) with the choice $\sigma_* = \sigma_0/4$, if the smallness condition on the energy¹⁰

$$R^2 < R_*^2 := \frac{2}{3(1+e)C_r}, \quad (27)$$

is satisfied, then the generating sequence \mathcal{X} defines an analytic canonical transformation on the domain $B_{\frac{2}{3}R}$ with the properties

$$B_{R/3} \subset T_{\mathcal{X}} B_{\frac{2}{3}R} \subset B_R \quad B_{R/3} \subset T_{\mathcal{X}}^{-1} B_{\frac{2}{3}R} \subset B_R.$$

Moreover, the deformation of the domain $B_{\frac{2}{3}R}$ is controlled by

$$z \in B_{\frac{2}{3}R} \quad \Rightarrow \quad \|T_{\mathcal{X}}(z) - z\| \leq 4^4 C_* R^3, \quad \|T_{\mathcal{X}}^{-1}(z) - z\| \leq 4^4 C_* R^3. \quad (28)$$

(iv) with the choice $\sigma_* = \sigma_0/4$, if (27) is satisfied, then the remainder is an analytic function on $B_{\frac{2}{3}R}$, and it is represented by a series of cyclically symmetric homogeneous polynomials $H_s^{(r)}$ of degree $2s + 2$

$$P^{(r+1)} = \sum_{s \geq r+1} H_s^{(r)} \quad H_s^{(r)} = \left(h_s^{(r)} \right)^\oplus, \quad (29)$$

and the seeds $h_s^{(r)}$ are of class $\mathcal{D}(2\tilde{C}_r^{s-1} C_{h_1}, \sigma_*)$ with $\tilde{C}_r = 96r^2 C_*$.

¹⁰Since R is the radius of the ball around the origin considered, the smallness condition is in total or specific energy depending on the phase space considered, i.e. respectively \mathcal{P}_2 or \mathcal{P}_∞ .

3.3 Some remarks

Some comments are in order. First and foremost we stress that our normal form Theorem holds both in a regime of small total energy and in a small *specific energy* regime. This fact is somewhat transparent in the Theorem's statement because the formulation is given in terms of small neighborhoods of the origin, the radius R being the small parameter: depending on the choice of the norm, euclidean or supremum one, the control is in total, respectively specific energy. From the technical viewpoint, this flexibility is embedded in Proposition 2.1 which is true both in \mathcal{P}_2 and in \mathcal{P}_∞ . From the point of view of the relevance of the result, the control with specific energy regime, joint with the uniformity in the number of degrees of freedom, give the validity of the normal form in the thermodynamic limit.

Clearly the validity of a normal form is only a first step: to fully exploit it, one has to give some precise control of the dynamics to ensure that, given suitable conditions on the initial datum, its evolution remains within the small neighborhood of the origin where the normal form holds. At present we are able to give such a control only in a regime of small *total energy*; indeed, in that case, a rather easy consequence of the normal form Theorem 3.1 is the almost invariance of H_Ω and \mathcal{Z} :

Corollary 3.1 *Let $z(0) \in B_{\frac{1}{9}R}$. There exists a constant C , independent of the main parameters R and a , such that the approximate integrals of motion H_Ω and \mathcal{Z} fulfill*

$$\begin{aligned} |H_\Omega(z(t)) - H_\Omega(z(0))| &\leq \Omega R^4, \\ |\mathcal{Z}(z(t)) - \mathcal{Z}(z(0))| &\leq R^4(C_{\zeta_0}\mu + C_{h_1}R^2), \end{aligned}$$

for times

$$|t| \leq \frac{C(1 - e^{-\sigma_*})^2}{C_{h_1}} (R^2 C_r)^{-r}. \quad (30)$$

In the time scale of the above Corollary, which is actually of the order $(Rr)^{-2r}$, one can think of the order r fixed, possibly at its maximal value of order $1/a$ according to (25), and then play with the small radius R , also provided it is satisfied the control $R \lesssim 1/r$ given by (27).

Actually the Corollary holds because H_Ω is equivalent to the euclidean norm, so that its conservation for long times is self-consistent: it comes from the structure of the normal form, and at the same time is enough to control the permanence in the right neighborhood of the origin. Unfortunately the control of the euclidean norm does not give a control of the sup norm.

In the small specific energy regime, in fact, we are still not able to exclude that an initial datum with the energy spread all over the chain could evolve into a localized state for which the sup norm would grow in a way essentially proportional to the number of degrees of freedom. And probably this could not be excluded at all. The results one could hope for, and which we are working on, are the following: either to show that such localization process takes a very long time, or that it happens for a set of initial data of small measure (both things asymptotically with the small parameter given by the specific energy).

Another kind of comments is related to the dependence of the smallness threshold R_* (defined in (27)) on the two different parameters involved in the perturbation construction: the coupling μ and the number of iteration steps r . We have:

- at fixed μ , R_* is monotonically decreasing with r (with a zero limit if one would be allowed to arbitrarily increase the number of steps r ; recall (25));
- at fixed $r \geq 1$, R_* increases when decreasing the coupling μ and it has an upper bound independent on μ .

One has to observe that, if we remove the coupling from the very beginning, i.e. $\mu \equiv 0$, the system is trivially composed of N identical anharmonic oscillators, and in such a case, it is known that the Birkhoff normal form procedure is defined on a ball of radius $0 < R_{**}(r) < 1$. Indeed our construction reduces to the standard one, once μ is set to zero, but our $R_*(r, \mu)$ does not converge¹¹ to $R_{**}(r)$ as $\mu \rightarrow 0$.

As a last comment we compare the present result with those of our previous works [13, 14]. There we constructed an (almost) conserved quantity, here we produce a normal form, which can give, in principle, much more information about the dynamics of the system. As a matter of fact, the application we sketch in the above Corollary 3.1 resemble very closely the results of the previous papers, with the following differences: here there's no need to exclude a small (with the Gibbs measure) set of initial data, but the result is valid only in total energy. In this sense it is somewhat complementary. But the above Corollary is only one of the possible applications once we have a normal form, which can shed more light on the structure of the Hamiltonian of the system. In the next Section we start to extract some information looking explicitly at the first step normal form, which turns out to be a generalized dNLS. We defer a deeper investigation in such a direction to forthcoming papers. We plan to explore possible applications of such a normal form: for example to the stability of Breathers (like in [19]) and MultiBreathers, or in order to give a justification for the otherwise formal use of the (G)dNLS to approximate the evolution of the KG model with small amplitude initial data. Moreover such a construction could be extended to the case of interactions, both linear and nonlinear, beyond the nearest neighbor: the scheme would be exactly the same, the first step being the study of the decay properties of the linear transformation (see Proposition 3.1), and the second one the control of the decay loss in the solution of the homological equation.

4 GdNLS model as normal form for the KG dynamics

Once the Hamiltonian is in the form (24), hence after the quadratic normalization, if we perform only one step of the perturbation scheme developed in Theorem 3.1, i.e. we choose $r = 1$ in (26), the transformed Hamiltonian reads

$$H^{(1)} = K + P^{(2)} , \quad K := H_\Omega + Z_0 + Z_1 , \quad P^{(2)} = \sum_{s \geq 2} H_s^{(1)} , \quad (31)$$

and the corresponding Hamilton equations are

$$\dot{z} = X_K(z) + X_{P^{(2)}}(z) . \quad (32)$$

4.1 The Generalized discrete Non Linear Schroedinger equation

In this part we want to stress and comment the fact that the simplified Hamiltonian K looks naturally as the Hamiltonian of a Generalized discrete Non Linear Schroedinger equation

¹¹Actually, since we have an upper bound on the number of steps r whenever $\mu \neq 0$, we were not interested in the optimization of all the estimates when $r \rightarrow \infty$.

(GdNLS). With the term *generalized* we mean that it includes interactions among sites which are also beyond the nearest-neighbors, both in the linear (Z_0) and in the nonlinear (Z_1) term.

We have indeed, by the normal form construction, the usual additional conserved quantity given by the ℓ^2 norm H_Ω

$$K = H_\Omega + \mathcal{Z} , \quad \mathcal{Z} = Z_0 + Z_1 , \quad \{H_\Omega, \mathcal{Z}\} = 0 . \quad (33)$$

Moreover, due to the decay property of the coefficients of such interactions, the Hamiltonian K turns out to be a perturbation of the dNLS model (here presented in real coordinates, see (46) for the standard one in complex coordinates)

$$H_{dNLS} = \sum_j \left[\frac{\Omega}{2}(q_j^2 + p_j^2) + \mu \frac{1}{2}(q_j q_{j+1} + p_j p_{j+1}) + \frac{3}{2}(q_j^2 + p_j^2)^2 \right] , \quad (34)$$

which is known to be a leading order normal form of the KG Hamiltonian, when the amplitude is taken proportional to $\sqrt{\mu}$, which means in the regime $E \sim \mu$, as discussed for example in the introduction of [20] (see also [9, 10] for other examples of the use of the dNLS in the study of a KG model).

Indeed, both the seeds ζ_0 and ζ_1 of the quadratic and quartic terms Z_0 and Z_1 include interactions which are exponentially small with the distance among sites, with the following expansions:

$$\zeta_0 = \sum_{m=1}^{[N/2]} \zeta_0^{(m)} , \quad \zeta_1 = \sum_{m=0}^{[N/2]} \zeta_1^{(m)} , \quad (35)$$

with supports for the components $\zeta_j^{(m)}$

$$S(\zeta_j^{(m)}) \subset [0, \dots, m] \cup [N-m, \dots, N] .$$

For the quadratic part we have an explicit expression:

$$\zeta_0^{(m)} = b_m [q_0(q_m + q_{N-m}) + p_0(p_m + p_{N-m})] , \quad |b_m| = \mathcal{O}(e^{-\sigma_0 m}) , \quad (36)$$

while for the quartic one we present here only a control of the norm of the components

$$\|\zeta_1^{(m)}\| \leq C'_{h_1} e^{-\sigma_0 m} = C'_{h_1} (2\mu)^m .$$

The effective computations of the monomials included in all the $\zeta_1^{(m)}$ is indeed a doable task, at least if supported by an algebraic manipulator program. We nevertheless defer such a task to future developments whenever it will be a necessary step.

With respect to the small parameter μ , the leading terms of ζ_0 and of ζ_1 are respectively the (resonant) nearest-neighbors interaction of the dNLS model and its nonlinear part (see, for comparison, formula (34), recalling that b_1 is $\mathcal{O}(\mu)$)

$$\zeta_0 = b_1 [q_0(q_1 + q_{N-1}) + p_0(p_1 + p_{N-1})] + \mathcal{O}(\mu^2) \quad \zeta_1 = \frac{3}{2}(q_0^2 + p_0^2)^2 + \mathcal{O}(\mu) .$$

Concerning ζ_1 we remark that, with respect to the expansion used in the normal form construction, we here¹² exploit the previously mentioned idea of taking the support $S(\zeta_1^{(m)})$

¹²The same approach could be extended elsewhere but it is beyond the purposes of the present work.

symmetrically centered around the 0-th variable. This provides the decay rate σ_0 , which represents a stronger condition than $\zeta_1 \in \mathcal{D}(C_{h_1}, \sigma_1)$, with $\sigma_1 < \sigma_0$, claimed in Proposition 3.1. This different expansion is a straightforward consequence of the following Lemma, whose proof is deferred to the Appendix

Lemma 4.1 *It is possible to select a seed h_1 such that*

$$h_1 = \sum_{m=0}^{[N/2]} h_1^{(m)}, \quad S(h_1^{(m)}) \subset [0, \dots, m] \cup [N - m, \dots, N]. \quad (37)$$

Moreover, there exists C'_{h_1} such that

$$\|h_1^{(m)}\| \leq C'_{h_1} e^{-m\sigma_0}.$$

4.2 Further comments on the construction of the normal form and on its relationship with the GdNLS model.

Here we aim at giving some comments on our results based on the remarks that we have two natural small parameters in our model, i.e. the coupling a and the energy E .

In the framework of a perturbation construction, at least at the formal level, the presence of a small parameter is usually exploited, by an expansion in its powers, to give a natural ordering of the terms which are dealt with at every step of the iterative procedure. If two small parameters are involved, we face a problem of gradation, which clearly comes from the lack of a natural ordering in \mathbb{Z}^2 . This is usually dealt with by choosing in advance a particular relation between the two parameters, i.e. by fixing a particular “regime”, thus effectively reducing the model to a system with a single perturbation parameter. For example, as already remarked, it is well known that by setting $E \sim a$ the standard dNLS arises as the first order normal form for the KG model (1) (see [20]). In such a case, it is ensured that, at least formally, at every step of the normal form construction, the remainder is smaller than the normal form terms.

Some comments are in order. The first is that for every different regime one aims to consider, a different normal form arises, in particular if higher order terms are involved: thus it is necessary to fix the ratio between the parameters in advance. Moreover, the procedure becomes rapidly quite cumbersome at the level of the selection of the terms which have to be dealt with at every step. An example is given in the next subsection 4.3, where the KG model is dealt with as described above, and with the choice of $a \lesssim E$, the dNLS appears after two steps.

In presence of more than one small parameter, even if one aims at producing a unique normal form to be used in different regimes of the parameters, the ordering used to treat the terms during the procedure must be chosen a priori, using some suitable criterion. In the present work, we deal with all orders in a – actually in a single step – for every fixed order in E . Indeed, every term Z_j in (26) contains the corrections for all the powers in a , and the index j relates to the order in E . This makes sense, and it is doable, exactly because the normalization with respect to a converges. The price one has to pay is that of course every term Z_j in (26) keeps also contributions which are smaller than those contained in the remainder: indeed, no matter how small E is, in every Z_j we have contributions containing a^l with l arbitrarily large, so that $E^j a^l < E^r$. The particular regime taken into account will determine how many terms, and which of them, are in this situation.

4.3 Comparison with a “standard” normalizing approach

Let us apply to the original Hamiltonian (1) the scaling $X_j = x_j/\sqrt{E}$, $Y_j = y_j/\sqrt{E}$; we obtain

$$H(X, Y, a, E) = \frac{1}{E} H(x(X), y(Y)) = \sum_{j=0}^{N-1} \left[\frac{X_j^2 + Y_j^2}{2} + a \frac{(X_{j+1} - X_j)^2}{2} + E \frac{X_j^4}{4} \right]; \quad (38)$$

the system actually presents two effective perturbation parameters, which are independent: E and a . In the limit $E \ll a < 1$ the dynamics is essentially governed by the whole quadratic part and the linear approximation prevails over the nonlinear effects. In the complementary case, $a \lesssim E < 1$, the on-site nonlinear dynamics is at least as relevant as the small coupling among nearby sites.

We are going to show that, if $a \lesssim E < 1$, then a “standard” normalizing procedure gives the dNLS as a resonant normal form of the KG model. Moreover, it essentially coincides¹³ with the leading part of the normal form (1) discussed in Section 4.

Birkhoff complex coordinates: We put the Hamiltonian into Birkhoff complex coordinates $\xi_j = (X_j + iY_j)/\sqrt{2}$, $i\eta_j = \bar{\xi}_j$, so that (38) reads

$$H(\xi, \eta) = h_\omega(\xi, \eta) + f^{(0)}(\xi, \eta, a) + H_1^{(0)}(\xi, \eta, E), \quad (39)$$

where the norm has been normalized to $h_\omega = 1$ and the “perturbation” is composed of

$$\begin{aligned} f^{(0)}(\xi, \eta, a) &= \frac{a}{4} \sum_{j=0}^{N-1} (\xi_{j+1}^2 + \xi_j^2 - \eta_{j+1}^2 - \eta_j^2 - 2\xi_{j+1}\xi_j + 2\eta_{j+1}\eta_j) + \\ &\quad + \frac{a}{2} \sum_{j=0}^{N-1} (\xi_{j+1} - \xi_j)(\bar{\xi}_{j+1} - \bar{\xi}_j), \\ H_1^{(0)}(\xi, \eta, E) &= \frac{E}{16} \sum_{j=1}^N (\xi_j^4 + \eta_j^4 + 4\xi_j^2|\xi_j|^2 - 4\eta_j^2|\xi_j|^2 + 6|\xi_j|^4). \end{aligned}$$

We perform a resonant normal form construction with respect to the resonant module

$$\mathcal{M}_\omega := \{k \in \mathbb{Z}^N | \langle k, \omega \rangle = 0\} = \{k \in \mathbb{Z}^N | k_1 + k_2 + \dots + k_N = 0\}.$$

First step: the first term $f^{(0)}$ is already split into a \mathcal{R}^2 and a \mathcal{N}^2 part: hence it is possible to define

$$Z_0^{(0)}(\xi, \eta) = \frac{a}{2} \sum_{j=1}^N |\xi_{j+1} - \xi_j|^2 = \Pi_{\mathcal{N}^2} f^{(0)}, \quad \{h_\omega, Z_0^{(0)}\} = 0, \quad (40)$$

as the a -resonant term and remove the Range part via a generating function $\mathcal{X}_0 = \mathcal{O}(a)$ which satisfies the usual homological equation

$$\{\mathcal{X}_0, h_\omega\} = Z_0^{(0)} - f^{(0)}, \quad \mathcal{X}_0 := L_{h_\omega}^{-1} [f^{(0)} - Z_0^{(0)}]. \quad (41)$$

¹³The only difference will be the coefficient of the next-neighbors coupling, since $b_1 \neq a/2$, although $b_1 = \mathcal{O}(a)$.

The change of coordinates $\mathcal{T}_{\mathcal{X}_0}$ gives the Hamiltonian the shape

$$H(\xi, \eta) = \sum_{j=1}^N |\xi_j^2| + \frac{a}{2} \sum_{j=1}^N |\xi_{j+1} - \xi_j|^2 + H_1^{(0)}(\xi, \eta) + \text{h.o.t.} , \quad (42)$$

where we have still used ξ, η to indicate the new coordinates. The higher order terms are

$$\text{h.o.t} = [\mathcal{T}_{\mathcal{X}_0} h_\omega - h_\omega - L_{\mathcal{X}_0} h_\omega] + [\mathcal{T}_{\mathcal{X}_0} f^{(0)} - f^{(0)}] + [\mathcal{T}_{\mathcal{X}_0} H_1^{(0)} - H_1^{(0)}] , \quad (43)$$

where the first two are of order $\mathcal{O}(a^2)$, while the last is of order $\mathcal{O}(aE)$.

The previous analysis forces to compare the newly generated quadratic term $f^{(1)} := \mathcal{T}_{\mathcal{X}_0} f^{(0)} - f^{(0)}$ with the quartic potential $H_1^{(0)}$, the first being $\mathcal{O}(a^2)$ and the second $\mathcal{O}(E)$. The part of the remainder which is $\mathcal{O}(aE)$ can be neglected at this step, being much smaller than $H_1^{(0)}$. If $a \lesssim E$ then $f^{(1)}$ can be transferred in the remainder and we can pass to consider $H_1^{(0)}$ as the next term to be normalized.

Second step with $a \lesssim E$: The resonant term in $H_1^{(0)}$ reads

$$Z_1^{(0)} = \frac{3E}{8} \sum_{j=0}^{N-1} |\xi_j|^4. \quad (44)$$

Thus removing through \mathcal{X}_1 all the Range terms of $H_1^{(0)}$, we obtain the normal form

$$H(\xi, \eta) = \overline{K} + \text{h.o.t.} , \quad \overline{K} := h_\omega + Z_0^{(0)} + Z_1^{(0)} , \quad \{h_\omega, \overline{K}\} = 0 . \quad (45)$$

Such a normal form of H is a dNLS model

$$\overline{K}(\xi, \eta) = \sum_{j=0}^{N-1} |\xi_j^2| + \frac{a}{2} \sum_{j=0}^{N-1} |\xi_{j+1} - \xi_j|^2 + \frac{3E}{8} \sum_{j=0}^{N-1} |\xi_j|^4 . \quad (46)$$

5 Proof of Theorem 3.1 and of Corollary 3.1

This section is devoted to the proof of Theorem 3.1 and of its immediate Corollary 3.1. In Section 5.1 we include a formal part, where we recall the process of construction of the normal form and discuss the solvability of the homological equation. In Section 5.2 we give all the quantitative estimates yielding to the statements of Theorem 3.1. Finally in Section 4.4 we prove Corollary 3.1.

5.1 Previous results: formal algorithm and solution of the homological equation

We recall here some basic facts on the formal algorithm we use to construct the normal form and to estimate its remainder: we refer to [12] for a detailed treatment.

Given a truncated generating sequence $\mathcal{X} = \{\mathcal{X}_s\}_{s=1, \dots, r}$, we define the linear operator $T_{\mathcal{X}}$ as

$$T_{\mathcal{X}} = \sum_{s \geq 0} E_s, \quad E_0 = \mathbb{I}, \quad E_s = \sum_{j=1}^s \binom{j}{s} L_{\mathcal{X}_j} E_{s-j} , \quad (47)$$

where $L_{\mathcal{X}_j} \cdot = \{\mathcal{X}_j, \cdot\}$ is the Lie derivative with respect to the flow generated by \mathcal{X}_j .

We look for \mathcal{X} and a function $H^{(r)}$ which represents a resonant *normal form* for the original Hamiltonian H , which means that $H^{(r)}$ satisfies the equation $T_{\mathcal{X}}H^{(r)} = H$, with $H^{(r)}$ of the form (26), where the normalized terms are in normal form in the sense that

$$\{H_{\Omega}, Z_s\} = 0 \quad \forall s = 0, \dots, r .$$

An immediate consequence is that H_{Ω} is an approximated first integral for the transformed Hamiltonian $H^{(r)}$, since

$$\{H_{\Omega}, H^{(r)}\} = \{H_{\Omega}, P^{(r+1)}\} .$$

We now translate the equation $T_{\mathcal{X}}Z = H$ into a formal recursive algorithm that allows us to construct both Z and \mathcal{X} . We take into account that our Hamiltonian has the particular form $H = H_0 + H_1$, where H_1 is a homogeneous polynomial of degree 4.

For $s \geq 1$ the generating function \mathcal{X}_s and the normalized term Z_s must satisfy the recursive set of *homological equations*

$$L_{H_0}\mathcal{X}_s = Z_s + \Psi_s; \quad (48)$$

where

$$\Psi_1 = H_1, \quad (49)$$

$$\Psi_s = \frac{s-1}{s} L_{\mathcal{X}_{s-1}} H_1 + \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j} Z_j, \quad s \geq 2 . \quad (50)$$

Our aim is to solve the homological equation (48) with the prescription that $L_{\Omega}Z_s = 0$ where $L_{\Omega} \cdot = \{H_{\Omega}, \cdot\}$ is the Lie derivative along the vector field generated by H_{Ω} as defined in (22). Thus we first point out the properties of the operator L_{Ω} , and then discuss how to solve the homological equation using a Neumann series expansion of the operator L_{H_0} , which is a μ perturbation of L_{Ω} .

5.1.1 The linear operator L_{Ω}

It is an easy matter to check that L_{Ω} maps the space of homogeneous polynomials into itself. It is also well known that L_{Ω} may be diagonalized via the canonical transformation

$$x_j = \frac{1}{\sqrt{2}}(\xi_j + i\eta_j), \quad y_j = \frac{i}{\sqrt{2}}(\xi_j - i\eta_j), \quad j = 1, \dots, N, \quad (51)$$

where $(\xi, \eta) \in \mathbb{C}^{2n}$ are complex variables. For a straightforward calculation gives

$$L_{\Omega}\xi^j\eta^k = i\Omega(|k| - |j|)\xi^j\eta^k, \quad (52)$$

where $|j| = |j_1| + \dots + |j_N|$ and similarly for $|k|$.¹⁴

Let us denote by $\mathcal{P}^{(s)}$ the linear space of the homogeneous polynomials of degree s in the $2n$ canonical variables $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n$. The kernel and the range of L_{Ω} are defined in the usual way, namely

$$\mathcal{N}^{(s)} = L_{\Omega}^{-1}(0), \quad \mathcal{R}^{(s)} = L_{\Omega}(\mathcal{P}^{(s)})$$

¹⁴ If $f(x, y) = \sum_{j,k} c_{j,k} x^j y^k$ is a real polynomial, then (51) produces a polynomial $g(\xi, \eta) = \sum_{j,k} b_{j,k} \xi^j \eta^k$ with complex coefficients $b_{j,k}$ satisfying $b_{j,k} = -b_{k,j}^*$, and conversely.

The property of L_Ω of being diagonal implies

$$\mathcal{N}^{(s)} \cap \mathcal{R}^{(s)} = \{0\} , \quad \mathcal{N}^{(s)} \oplus \mathcal{R}^{(s)} = \mathcal{P}^{(s)} .$$

Thus the inverse $L_\Omega^{-1} : \mathcal{R}^{(s)} \rightarrow \mathcal{R}^{(s)}$ is uniquely defined.

5.1.2 The linear operator L_{H_0}

We come now to the solution of the homological equation (48). In view of (22) we have $L_{H_0} = L_\Omega + L_{Z_0}$, or equivalently $L_{H_0} = L_\Omega(\mathbb{I} + L_\Omega^{-1}L_{Z_0})$. Thus we have

$$L_{H_0}^{-1} = (\mathbb{I} + K)^{-1}L_\Omega^{-1} , \quad K := L_\Omega^{-1}L_{Z_0} , \quad (53)$$

and using the Neumann's series we formally get $(\mathbb{I} + K)^{-1} = \sum_{l \geq 0} (-1)^l K^l$.

A general consideration is the following. Let us consider L_{H_0} on a topological space $\mathcal{P}^{(s)}$ endowed with any norm $\|\cdot\|$. The next Proposition (see again [14], Section 4) claims that, although we ignore its Kernel and Range, we can invert L_{H_0} on $\mathcal{R}^{(s)}$.

Proposition 5.1 (see [14]) *If the restriction of K to $\mathcal{R}^{(s)}$ satisfies*

$$\|K\|_{\text{op}} < 1, \quad (54)$$

then for any $g \in \mathcal{R}^{(s)}$, there exists a unique element $f \in \mathcal{R}^{(s)}$ of the form

$$f = \sum_{l \geq 0} (-1)^l K^l g , \quad \text{such that} \quad (\mathbb{I} + K)f = g .$$

5.2 Proof of Theorem 3.1

In order to show that the exponential decay of interactions is preserved by our construction, our first aim (as in [14]) is to show that the functions \mathcal{X}_s , Ψ_s and Z_s , that are generated by the formal construction, are of class $\mathcal{D}(\cdot, \sigma_s)$, with suitable values of σ_s and with some constant to be evaluated in place of the dot. A second step pertains the estimate of the remainder $\mathcal{R}^{(r+1)}$, which has still to be expressed in terms of \mathcal{X}_s , Ψ_s and Z_s . We conclude with the estimates of the Hamiltonian vector fields of the generating functions \mathcal{X}_s , which allow to control the deformation of the transformation $\mathcal{T}_\mathcal{X}$.

Let us pick $\sigma_* < \sigma_1$, and recall we have defined in Theorem 3.1

$$\sigma_j := \sigma_1 - \frac{(j-1)}{r}(\sigma_1 - \sigma_*) , \quad \text{for } j = 1, \dots, r , \quad (55)$$

so that $\sigma_1 > \dots > \sigma_r > \sigma_{r+1} = \sigma_*$. We shall repeatedly use the following elementary estimates. By the general inequality

$$1 - e^{-x} \geq \left(\frac{x}{a}\right)(1 - e^{-a}) \quad \text{for } 0 \leq x \leq a , \quad (56)$$

for $0 \leq j < s \leq r$ we get

$$\begin{aligned} 1 - e^{-\max(\sigma_j, \sigma_{s-j})} &\geq \frac{1 - e^{-\sigma_0}}{\sigma_0} \max(\sigma_j, \sigma_{s-j}) \geq \frac{(1 - e^{-\sigma_0})}{4} , \\ 1 - e^{-(\sigma_j - \sigma_s)} &\geq \frac{s-j}{r} (1 - e^{-(\sigma_0 - \sigma_*)}) \left[\frac{\min(\sigma_0 - \sigma_1, \sigma_1 - \sigma_*)}{\sigma_0 - \sigma_*} \right] . \end{aligned} \quad (57)$$

To get the first inequality we make use of $\max(\sigma_j, \sigma_{s-j}) \geq (\sigma_1 + \sigma_*)/2 \geq \sigma_*$ and of the hypothesis $\sigma_* \geq \sigma_0/4$. Concerning the second of (57), take first $0 = j < s \leq r$ and apply (56) to get

$$\begin{aligned} 1 - e^{-(\sigma_0 - \sigma_s)} &\geq \left[\frac{(\sigma_0 - \sigma_1) + \frac{s-1}{r}(\sigma_1 - \sigma_*)}{\sigma_0 - \sigma_*} \right] (1 - e^{-(\sigma_0 - \sigma_*)}) > \\ &> \frac{s}{r} \left(\frac{\sigma_0 - \sigma_1}{\sigma_0 - \sigma_*} \right) (1 - e^{-(\sigma_0 - \sigma_*)}) ; \end{aligned}$$

then take $1 \leq j < s \leq r$ and apply (56) to get

$$1 - e^{-(\sigma_j - \sigma_s)} \geq \left(\frac{\sigma_j - \sigma_s}{\sigma_0 - \sigma_*} \right) (1 - e^{-(\sigma_0 - \sigma_*)}) = \frac{s-j}{r} \left(\frac{\sigma_1 - \sigma_*}{\sigma_0 - \sigma_*} \right) (1 - e^{-(\sigma_0 - \sigma_*)}) .$$

Estimate of the homological equation. We summarize here the main result and comments which can be found in [14], proof included, about the estimate of the homological equations (48).

In order to proceed we first define

$$E_{0*} := \frac{\min(\sigma_0 - \sigma_1, \sigma_1 - \sigma_*)}{\sigma_0 - \sigma_*} . \quad (58)$$

and then we give consistent values for the constants of Theorem 3.1

$$\begin{aligned} \mu_* &= \frac{\Omega(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})E_{0*}}{8C_{\zeta_0}e^{\sigma_1}} , \\ \gamma &= 2\Omega \left(1 - \frac{r\mu}{\mu_*} \right) , \\ C_* &= \frac{C_{h_1}}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})E_{0*}} . \end{aligned} \quad (59)$$

Lemma 5.1 *Let $G = g^\oplus \in \mathcal{R}^{(2s+2)}$ be a cyclically symmetric homogeneous polynomial of degree $2s+2$ of class $\mathcal{D}(C_g, \sigma_s)$. Let K as defined in (53) and assume*

$$C_K := \frac{4C_0e^{-(\sigma_0 - \sigma_1)}}{\Omega(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})E_{0*}} \leq \frac{1}{2r} . \quad (60)$$

Then there exists a cyclically symmetric homogeneous polynomial $\mathcal{X} = \chi^\oplus \in \mathcal{R}^{(2s+2)}$ which solves $L_{H_0}\mathcal{X} = G$; moreover χ is of class $\mathcal{D}(C_g/\gamma, \sigma_s)$ with

$$\gamma = 2\Omega(1 - rC_K) . \quad (61)$$

Remark 5.1 *In Proposition 5.1 we ask $\|K\|_{\text{op}} < 1$ to simply perform the inversion. In the above Lemma 5.1, condition (60) reads as $\|K\|_{\text{op}} < 1/2$, and this stronger requirement is to control the small divisors (61).*

We emphasize that in view of the first of (59) we have $C_K = \mu/\mu_*$. Therefore, condition (60) reads $2r\mu < \mu_*$, which is the smallness condition for μ of Theorem 3.1. Furthermore this gives the value of γ in (59). Moreover, the constant γ is evaluated as independent of s , but seems to depend on the degree r of truncation of the first integral. However, in view of the condition on μ we have $\Omega \leq \gamma \leq 2\Omega$.

Having thus proved that the homological equation can be solved, the statement (i) of Theorem 3.1 follows.

Iterative estimates on the generating sequence. We follow the same procedure used in [14], Subsection 4.2. We recall that the generating sequence is found by recursively solving the homological equations $L_{H_0}\chi_s = Z_s + \Psi_s$ for $s = 1, \dots, r$ with

$$\begin{aligned}\Psi_1 &= H_1, \\ \Psi_s &= \left(\frac{s-1}{s}\right)L_{\mathcal{X}_{s-1}}H_1 + \sum_{l=1}^{s-1} \left(\frac{l}{s}\right)E_{s-l}Z_l, \\ E_s Z_l &= \sum_{j=1}^s \left(\frac{j}{s}\right)L_{\mathcal{X}_j}E_{s-j}Z_l \quad \text{for } s \geq 1.\end{aligned}\tag{62}$$

Our aim is to find positive constants $C_{\psi,1}, \dots, C_{\psi,r}$ so that Ψ_s is of class $\mathcal{D}(C_{\psi,s}, \sigma_s)$. In view of lemma 5.1 this implies that Z_s is of class $\mathcal{D}(C_{\zeta,s}, \sigma_s)$ with $C_{\zeta,s} = C_{\psi,s}$ and χ_s is of class $\mathcal{D}(C_{\chi,s}, \sigma_s)$ with $C_{\chi,s} = C_{\psi,s}/\gamma$. Meanwhile we also find constants $C_{\zeta,s,l}$ such that $E_s Z_l$ is of class $\mathcal{D}(C_{\zeta,s,l}, \sigma_{s+l})$ whenever $s+l \leq r$.

We look for a constant B_r and two sequences $\{\eta_s\}_{1 \leq s \leq r}$ and $\{\theta_s\}_{1 \leq s \leq r}$ such that

$$\begin{aligned}C_{\psi,1} &\leq \eta_1 C_{h_1}, \quad C_{\zeta,0,1} \leq \eta_1 \theta_0 C_{h_1}, \\ C_{\psi,s} &\leq \frac{\eta_s}{s} B_r^{s-1} C_{h_1} \quad \text{for } s > 1, \\ C_{\zeta,s,l} &\leq \theta_s \eta_l B_r^{s+l-1} C_{h_1} \quad \text{for } s \geq 1, l \geq 1.\end{aligned}\tag{63}$$

In view of $\Psi_1 = H_1$ and of $E_0 Z_1 = Z_1$ we can choose $\eta_1 = \theta_0 = 1$. By (62) and using lemmas 5.1 and 6.2 together with corollary 6.1 we get the recursive relations

$$\begin{aligned}C_{\zeta,s,l} &\leq \frac{4}{s} \sum_{j=1}^{s-1} \frac{j(s+l-j)\eta_j \eta_l \theta_{s-j}}{(1 - e^{-\max(\sigma_j, \sigma_{s+l-j}) + \sigma_{s+l}})(1 - e^{-\max(\sigma_j, \sigma_{s+l-j})})} \cdot \frac{B_r^{s+l-2} C_{h_1}^2}{\gamma}, \\ C_{\psi,s} &\leq \left(\frac{8(s-1)\eta_{s-1} C_{h_1}}{s(1 - e^{-(\sigma_0 - \sigma_s)})(1 - e^{-\sigma_0})} + \sum_{l=1}^{s-1} \frac{l B_r}{s} \eta_l \theta_{s-l} \right) \cdot \frac{B_r^{s-2} C_{h_1}}{\gamma},\end{aligned}\tag{64}$$

Using the first of (57), with $s+l$ instead of s , we have

$$1 - e^{-\max(\sigma_j, \sigma_{s+l-j})} > \frac{1 - e^{-\sigma_0}}{4}.$$

Using the second of (57) in a similar way to deal with

$$1 - e^{-[\max(\sigma_j, \sigma_{s+l-j}) - \sigma_{s+l}]} \geq \frac{s+l - \min(j, s+l-j)}{r} \left(1 - e^{-(\sigma_0 - \sigma_*)}\right) E_{0*},$$

and setting

$$B_r = \frac{16 C_{h_1} r}{\gamma(1 - e^{-(\sigma_0 - \sigma_*)})(1 - e^{-\sigma_0}) E_{0*}}.\tag{65}$$

we get

$$\begin{aligned}C_{\zeta,l,s} &\leq \frac{1}{s} \sum_{j=1}^{s-1} j \eta_j \eta_l \theta_{s-j} B_r^{s+l-1} C_{h_1}, \\ C_{\psi,s} &\leq \left(\frac{1}{s} \eta_{s-1} + \sum_{l=1}^{s-1} \frac{l}{s} \eta_l \theta_{s-l} \right) B_r^{s-1} C_{h_1}.\end{aligned}\tag{66}$$

Thus the inequalities (63) are satisfied by the sequences recursively defined as

$$\begin{aligned}\theta_s &:= \sum_{j=1}^{s-1} \frac{j}{s} \eta_j \theta_{s-j} & \text{for } s \geq 1, \\ \eta_s &:= \eta_{s-1} + \sum_{j=1}^{s-1} j \eta_j \theta_{s-j} & \text{for } s \geq 2.\end{aligned}$$

starting with $\eta_1 = \theta_0 = 1$. It is possible, recalling also that $s \leq r$, to show that (see [14])

$$\eta_s < 4^{s-1} r^{s-1}.$$

Replacing this and (64) in the inequality (63) for $C_{\psi,s}$ and recalling that $C_{\chi,s} \leq C_{\psi,s}/\gamma$ we have

$$C_{\chi,s} \leq (64r^2 C_*)^{s-1} \frac{C_{h_1}}{\gamma s}, \quad C_* = \frac{C_{h_1}}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})E_{0*}}.$$

This proves the statement (ii) of Theorem 3.1 with the estimated value of C_* in (59). The statement (iii) also follows in view of $C_{\zeta,s} \leq C_{\psi,s}$.

The remainder of the normal form. We combine here the formal algorithm developed in Chapter 4 of [12] with the previous estimates on χ_s, Ψ_s, Z_s . In general, the remainder $P^{(r+1)}$ can be written using the operators D_s

$$D_s := - \sum_{j=1}^s \binom{j}{s} D_{s-j} L_{\chi_j} \quad D_0 := \mathbb{I}, \quad (67)$$

which define the inverse of the canonical transformation generated by \mathcal{T}_χ

$$\mathcal{T}_\chi^{-1} := \sum_{s \geq 0} D_s.$$

Thus, as stressed in Paragraph 5.2.3 of [12], one has

$$P^{(r+1)} = \sum_{s > r} H_s^{(r)}, \quad H_s^{(r)} := \sum_{j=0}^{s-1} D_j H_{s-j} - \sum_{j=1}^s \frac{j}{s} D_{s-j} (\Psi_j - Z_j). \quad (68)$$

In our specific case

$$H_s^{(r)} = D_{s-1} H_1 - \sum_{j=1}^r \frac{j}{s} D_{s-j} (\Psi_j - Z_j);$$

indeed the Hamiltonian has initially only one nonlinear term, namely a quartic polynomial H_1 , and $\chi_s = 0$ for all $s \geq r+1$.

Let us consider a generic cyclically symmetric polynomial $F \in \mathcal{D}(C_f, \sigma_h)$ of degree $2h+2$ with $h = 1, \dots, r$. We are interested in estimating $D_l F$; more precisely, we look for a sequence d_l such that $D_l F \in \mathcal{D}(d_l C_f, \sigma_*)$. We notice that, since $\sigma_* < \sigma_r$, it is possible to deal with $L_{\chi_r}(\Psi_r - Z_r)$, where both functions in the Poisson brackets belong to $\mathcal{D}(\cdot, \sigma_r)$. For any $1 \leq j \leq r$, one has $L_{\chi_j} F \in \mathcal{D}(c_j, \sigma_*)$ for a some suitable c_j . We estimate c_j using (78)

$$c_j \leq \frac{4(j+1)(h+1)C_{\chi_j}C_f}{(1 - e^{-\max(\sigma_j, \sigma_h)})(1 - e^{-\max(\sigma_j, \sigma_h) + \sigma_*})}.$$

From the first of (57) we have

$$\frac{1}{1 - e^{-\max(\sigma_j, \sigma_h)}} \leq \frac{4}{1 - e^{-\sigma_0}} ,$$

and

$$\begin{aligned} 1 - e^{-\max(\sigma_j, \sigma_h) + \sigma_*} &\geq \left(\frac{\max(\sigma_j, \sigma_h) - \sigma_*}{\sigma_0 - \sigma_*} \right) \left(1 - e^{-(\sigma_0 - \sigma_*)} \right) \geq \\ &\geq \left(\frac{\sigma_r - \sigma_*}{\sigma_0 - \sigma_*} \right) \left(1 - e^{-(\sigma_0 - \sigma_*)} \right) = \left(\frac{\sigma_1 - \sigma_*}{\sigma_0 - \sigma_*} \right) \left(\frac{1 - e^{-(\sigma_0 - \sigma_*)}}{r} \right) \end{aligned}$$

hence by inserting $\sigma_1 = \sigma_0/2$ (see (23)) and making use of $\sigma_* = \sigma_0/4$ (see additional hypothesis in claims (iv) and (v) of Theorem 3.1) we obtain

$$\frac{1}{1 - e^{-\max(\sigma_j, \sigma_h) + \sigma_*}} \leq \frac{3r}{1 - e^{-(\sigma_0 - \sigma_*)}} .$$

By combining the two above estimates we get

$$c_j \leq \frac{48(j+1)(h+1)(r+1)}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} C_{\chi_j} C_f .$$

If we use¹⁵

$$C_{\chi_j} \leq C_r^{j-1} \frac{C_{h_1}}{\gamma j} \quad C_r = \frac{192r^2 C_{h_1}}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} ,$$

then, using $h+1 \leq r+1$, we have

$$c_j \leq C_r^{j-1} \frac{192r^2 C_{h_1}}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} C_f = C_r^j C_f . \quad (69)$$

From (67), the sequence d_l has to satisfy

$$d_l C_f \leq \sum_{j=1}^l \frac{j}{l} d_{l-j} c_j \leq \sum_{j=1}^l \frac{j}{l} d_{l-j} C_r^j C_f .$$

We look for d_l of the form $d_m := \theta_m C_r^m$. By inserting in the above recursive inequality and removing the common factor C_r^l we obtain for θ_l the relation¹⁶

$$\theta_l \leq \sum_{j=1}^l \frac{j}{l} \theta_{l-j} < \sum_{j=0}^{l-1} \theta_j \leq 2^l ,$$

thus we have

$$D_l F \in \mathcal{D}(\tilde{C}_r^l C_f, \sigma_*) , \quad \tilde{C}_r := 2C_r . \quad (70)$$

We apply the above result to the elements giving $H_s^{(r)}$; for the sake of brevity we define $F_j := \Psi_j - Z_j \in \mathcal{D}\left(\frac{2C_{h_1}}{j} C_r^{j-1}, \sigma_*\right)$, so we have, for all $s \geq r+1$,

$$\begin{aligned} D_{s-1} H_1 &\in \mathcal{D}\left(C_{h_1} \tilde{C}_r^{s-1}, \sigma_*\right) \\ D_{s-j} F_j &\in \mathcal{D}\left(\frac{4C_{h_1}}{j2^j} \tilde{C}_r^{s-1}, \sigma_*\right) \implies C_{H_s^{(r)}} \leq \tilde{C}_r^{s-1} C_{h_1} \left[1 + \sum_{j=1}^r \frac{4}{s2^j} \right] < 2\tilde{C}_r^{s-1} C_{h_1} . \end{aligned}$$

¹⁵Since we are making use of $\sigma_* = \sigma_0/4$, it is $E_{0*} = 1/3$.

¹⁶ θ_l is dominated by $x_l := \sum_{j=0}^{l-1} x_j$, which gives $x_l = 2^l$ once it is set $x_0 = 1$.

The canonical transformation. In order to control the domain of validity of the normal form, we need to estimate the deformation due to the canonical transformation $\mathcal{T}_{\mathcal{X}}$. Since it is obtained by composition of r consecutive normal form steps $\mathcal{T}_r := \mathcal{T}_{\mathcal{X}_r}$, we set

$$d = r\delta_1 \quad \delta_s = \delta_1 = \frac{d}{r} \quad \forall s = 1, \dots, r,$$

and we require for any $s = 1, \dots, r$ the generic transformation \mathcal{T}_s to map

$$\mathcal{T}_s : B_{R_s} \rightarrow B_{R_{s-1}}, \quad \text{where } R_s := R_{s-1} - \delta_s, \quad R_0 \equiv R.$$

Hence, by composition, the whole transformation \mathcal{T} maps B_{R-d} to B_R . We have

Lemma 5.2 *Let $\delta_s < R_{s-1}$ and $\|X_{\mathcal{X}_s}\|_{R_{s-1}}^{\oplus} < \delta_s/e$, then for $|t| \leq 1$ and $\|z\| \leq R_s$ one has*

$$\|\mathcal{T}_s^t(z) - (z)\| \leq \frac{1}{1 - c_s} \|X_{\mathcal{X}_s}\|_1^{\oplus} \|z\|^{2s+1}, \quad \text{with } c_s := \frac{e}{\delta_s} \|X_{\mathcal{X}_s}\|_{R_{s-1}}^{\oplus}. \quad (71)$$

proof: By series expansion we have

$$\mathcal{T}_s^t(z) - z = \sum_{p \geq 1} \frac{1}{p!} L_{\mathcal{X}_s}^{p-1} X_{\mathcal{X}_s}(z) = \sum_{p \geq 0} \frac{1}{(p+1)!} L_{\mathcal{X}_s}^p X_{\mathcal{X}_s}(z).$$

By applying Corollary 6.3 (with both \mathcal{X} and F of the Corollary equal to \mathcal{X}_s)

$$\|L_{\mathcal{X}_s}^p X_{\mathcal{X}_s}(z)\| \leq \prod_{j=0}^{p-1} [(j+1)s - j] \left(\|X_{\mathcal{X}_s}\|_1^{\oplus} \right)^{p+1} \|z\|^s \|z\|^{p(s-1)}. \quad (72)$$

Let us take initially the case $p = 2$ with $\|z\| \leq R - 2d$ and $d = \delta_s/2$; we apply twice (86) and (85)

$$(2s-1)s(R-2d)^{2s-2} \leq \frac{1}{d}s(R-d)^{2s-1} \leq \frac{1}{d}R^s s(R-d)^{s-1} \leq \frac{1}{d^2}R^{2s}.$$

We define $d = \delta_s/p$ and by iteration of the same argument we get

$$\prod_{j=0}^{p-1} [(j+1)s - j](R - pd)^{ps-p} \leq \frac{1}{d^p} R^{ps} = \frac{p^p}{\delta_s^p} R^{ps} \leq p! \left(\frac{e}{\delta_s} \right)^p (R^s)^p. \quad (73)$$

We can estimate (72) as follows

$$\|L_{\mathcal{X}_s}^p X_{\mathcal{X}_s}(z)\| \leq p! \|X_{\mathcal{X}_s}\|_1^{\oplus} \|z\|^s \left(\frac{e \|X_{\mathcal{X}_s}\|_R^{\oplus}}{\delta_s} \right)^p.$$

Thus for $\|z\| \leq R - \delta_s$ we have

$$\left\| \sum_{p \geq 0} \frac{1}{(p+1)!} L_{\mathcal{X}_s}^p X_{\mathcal{X}_s}(z) \right\| \leq \|X_{\mathcal{X}_s}\|_1 \|z\|^s \sum_{p \geq 0} \left(\frac{e}{\delta_s} \|X_{\mathcal{X}_s}\|_R^{\oplus} \right)^p,$$

which gives the claim. \square

Corollary 5.1 *Let $\delta_s < R_{s-1}$ and also assume*

$$\left\| X_{\mathcal{X}_s} \right\|_{R_{s-1}}^{\oplus} < \frac{\delta_s}{1+e} , \quad (74)$$

then, for $z \in B_{R_s}$, one has

$$\| \mathcal{T}_s(z) - z \| \leq (1+e) \left\| X_{\mathcal{X}_s} \right\|_{R_{s-1}}^{\oplus} \leq \delta_s \quad \Rightarrow \quad \mathcal{T}_s : B_{R_s} \rightarrow B_{R_{s-1}} . \quad (75)$$

proof: Given the control of z (i.e. $\|z\| \leq R_s < R_{s-1}$), and the degree of $X_{\mathcal{X}_s}$, one has

$$\left\| X_{\mathcal{X}_s} \right\|_1^{\oplus} \|z\|^{2s+1} < \left\| X_{\mathcal{X}_s} \right\|_{R_{s-1}}^{\oplus} ;$$

by the definition of c_s in (71) and by hypothesis (74), one has $c_s \leq \frac{e}{e+1}$, i.e. $\frac{1}{1-c_s} \leq 1+e$; inserting these inequalities into the estimate of (71) one gets the thesis. \square

Conclusion of the proof. We now make use of Corollary 5.1 to conclude the proof of (iv). From (18) we have

$$\left\| X_{\mathcal{X}_s} \right\|_{R_{s-1}}^{\oplus} \leq \left\| X_{\mathcal{X}_s} \right\|_R^{\oplus} \leq \frac{4(s+1)R^{2s+1}C_{\mathcal{X}_s}}{(1-e^{-\sigma_s})^2} ;$$

we apply (56)

$$\begin{aligned} 1 - e^{-\sigma_s} &> \left(\frac{\sigma_s}{\sigma_0 - \sigma_s} \right) (1 - e^{-(\sigma_0 - \sigma_*)}) > \left(\frac{\sigma_*}{\sigma_0 - \sigma_*} \right) (1 - e^{-(\sigma_0 - \sigma_*)}) \\ 1 - e^{-\sigma_s} &> \left(\frac{\sigma_s}{\sigma_0} \right) (1 - e^{-\sigma_0}) > \left(\frac{\sigma_*}{\sigma_0} \right) (1 - e^{-\sigma_0}) \end{aligned}$$

and inserting $\sigma_* = \sigma_0/4$ we get

$$\frac{1}{(1 - e^{-\sigma_s})^2} < \frac{\sigma_0(\sigma_0 - \sigma_*)}{\sigma_*^2(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} = \frac{12}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} ,$$

thus the smallness condition can be replaced by

$$\frac{48(s+1)R^{2s+1}C_{\mathcal{X}_s}}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} < \frac{\delta_s}{1+e} .$$

We recall that

$$C_{\mathcal{X}_s} = \frac{1}{s\gamma} C_{h_1} C_r^{s-1} \quad C_r = \frac{192r^2 C_{h_1}}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} ,$$

so that the field $X_{\mathcal{X}_s}$ fulfills

$$\left\| X_{\mathcal{X}_s} \right\|_{R_{s-1}}^{\oplus} \leq \frac{48(s+1)R^{2s+1}C_{\mathcal{X}_s}}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} \leq \frac{R}{2r^2} a_s , \quad \text{with } a_s := (R^2 C_r)^s . \quad (76)$$

Under the smallness condition (27), the sequence a_s is controlled by a geometrically decreasing one, $a_s < \left(\frac{2}{3(1+e)} \right)^s$, so that we can think at \mathcal{T}_s as a sequence of increasingly smaller deformation of the identity, the first \mathcal{T}_1 being the biggest.

With the choice $d := \frac{R}{3}$, condition (74) is ensured by imposing

$$\frac{R}{2r^2}a_1 < \frac{\delta_s}{1+e} = \frac{\delta_1}{1+e} = \frac{d}{r(1+e)} = \frac{R}{3r(1+e)},$$

which is fulfilled provided $a_1 < \frac{2}{3(1+e)}$, which in turn is (27).

To conclude the proof of Theorem 3.1 we still have to prove the smallness of the deformation of the domain $B_{\frac{2}{3}R}$. This is obtained summing up all the (geometrically decreasing) deformations (75) of the iteratively defined domains B_s . Indeed, we exploit (76) to collect all the r deformations

$$\|T\chi(z) - z\| \leq \frac{(1+e)R}{2r^2} \left[\sum_{s=1}^r a_s \right] \leq \frac{(1+e)R^3 C_r}{r^2} < R^3 4^4 C_*.$$

5.3 Proof of Corollary 3.1

Since by hypothesis $z(0) \in B_{\frac{R}{9}}$, the transformed initial datum lies in $\tilde{z}(0) \in B_{\frac{4}{9}R}$. Let τ the escape time from the ball of radius $\frac{2}{3}R$, then for all $|t| < \tau$ it holds

$$\begin{aligned} |H_\Omega(\tilde{z}(t)) - H_\Omega(\tilde{z}(0))| &\leq \int_0^t |\{H_\Omega, P^{(r+1)}\}\tilde{z}(s)| ds \leq \|X_{H_\Omega}\|_{\frac{2}{3}R} \|X_{P^{(r+1)}}\|_{\frac{2}{3}R} |t| = \\ &= \frac{2}{3} \Omega R \|X_{P^{(r+1)}}\|_{\frac{2}{3}R} |t|. \end{aligned}$$

In order to deal with the remainder, we apply (29)

$$P^{(r+1)} = \sum_{s \geq r+1} H_s^{(r)}, \quad \left\| X_{H_s^{(r)}} \right\|_{\frac{2}{3}R} = \left\| X_{H_s^{(r)}} \right\|_1^\oplus \left(\frac{2}{3} R \right)^{2s+1};$$

from Theorem 3.1, (v), and Lemma 2.4 we know that

$$\left\| X_{H_s^{(r)}} \right\|_1^\oplus \leq \frac{8(s+1)}{(1-e^{-\sigma_*})^2} C_{h_s^{(r)}},$$

hence the Hamiltonian field of the whole remainder $P^{(r+1)}$ fulfills

$$\|X_{P^{(r+1)}}\|_{\frac{2}{3}R} \leq \frac{16C_{h_1}}{(1-e^{-\sigma_*})^2} R^{2r+3} \tilde{C}_r^r \left[\sum_{h \geq 0} (h+3) \left(\frac{2}{3} R^2 \tilde{C}_r \right)^h \right].$$

From the smallness assumption (27) we have $R^2 \tilde{C}_r < 1/(1+e)$, then

$$|H_\Omega(\tilde{z}(t)) - H_\Omega(\tilde{z}(0))| \leq \frac{2^5 \Omega C_{h_1}}{(1-e^{-\sigma_*})^2} \left(\frac{2}{3} R \right)^{2r+4} \tilde{C}_r^r |t|,$$

which gives for a suitable C

$$|H_\Omega(\tilde{z}(t)) - H_\Omega(\tilde{z}(0))| < \Omega R^4, \quad |t| \leq \frac{C(1-e^{-\sigma_*})^2}{C_{h_1}} \left(\frac{2}{3} R^2 \tilde{C}_r \right)^{-r}.$$

The variation in the original coordinates follows from two facts. The first and main one is that by controlling H_Ω we are controlling the escape time from the ball where we started. The second one is the deformation of the canonical transformation $T_\mathcal{X}$, which according to (28) gives

$$|H_\Omega(\tilde{z}) - H_\Omega(z)| \leq \Omega R^4 .$$

We proceed in the same way to control the variation of \mathcal{Z} . From its definition in (26)

$$\|X_\mathcal{Z}\|_{\frac{2}{3}R}^\oplus \leq \sum_{s=0}^r \|X_{Z_s}\|_{\frac{2}{3}R}^\oplus ;$$

we make use of Lemma 2.4 to get

$$\|X_{Z_0}\|_1^\oplus \leq \frac{8C_{\zeta_0}\mu}{(1 - e^{-\sigma_*})^2} , \quad \|X_{Z_s}\|_1^\oplus \leq \frac{8C_{h_1}C_r^{s-1}}{(1 - e^{-\sigma_*})^2} \quad s = 1, \dots, r .$$

where the factor μ in the first estimate follows from the fact that $\zeta_0^{(0)} = 0$. □

6 Appendix

6.1 Poisson brackets of cyclically symmetric polynomials

The following Lemma produces a general estimate of the Poisson bracket specially adapted to the case of cyclically symmetric polynomials. It is crucial for the control of the dependence on N of the norms of “extensive” functions generated by our perturbation scheme. Its proof, and those of the subsequent statements, are collected in the Appendix of [14].

Lemma 6.1 (see [14]) *Let $f(x, y)$ and $g(x, y)$ be homogeneous polynomials respectively of degree r and s . Then $\{f, g\}$ is a homogeneous polynomial of degree $r + s - 2$, and one has*

$$\|\{f, g\}\|_1 \leq rs \|f\|_1 \|g\|_1 .$$

Moreover, there exists a seed of $\{f^\oplus, g^\oplus\}$ such that one has

$$\|\{f^\oplus, g^\oplus\}\|_1^\oplus \leq rs \|f^\oplus\|_1^\oplus \|g^\oplus\|_1^\oplus . \quad (77)$$

The next statements provide the basic estimates for controlling the exponential decay.

Lemma 6.2 (see [14]) *Let F, G be cyclically symmetric homogeneous polynomials of degree r', r'' respectively. Let the seeds f, g be of class $\mathcal{D}(C_f, \sigma')$ and $\mathcal{D}(C_g, \sigma'')$, respectively, and let $\sigma < \min(\sigma', \sigma'')$. Then there exists $C_h \geq 0$ such that the seed h of $H = \{F, G\}$ is of class $\mathcal{D}(C_h, \sigma)$. An explicit estimate is*

$$C_h = \frac{r'r''C_fC_g}{(1 - e^{-\max(\sigma', \sigma'')})(1 - e^{-\max(\sigma', \sigma'') + \sigma})} . \quad (78)$$

Corollary 6.1 (see [14]) *If in lemma 6.2 we have $\sigma' \neq \sigma''$ then we may set $\sigma = \min(\sigma', \sigma'')$ and*

$$C_h = \frac{r'r''C_fC_g}{(1 - e^{-\max(\sigma', \sigma'')})(1 - e^{-|\sigma' - \sigma''|})} .$$

Corollary 6.2 (see [14]) *If in lemma 6.2 we have $\sigma' > \sigma''$ and $f^{(0)} = 0$, i.e., $f = \sum_{m \geq 1} f^{(m)} = O(e^{-\sigma'})$ then we may set $\sigma = \sigma''$ and*

$$C_h = \frac{2e^{-(\sigma' - \sigma'')} r' r'' C_f C_g}{(1 - e^{-\sigma'}) (1 - e^{-(\sigma' - \sigma'')})} . \quad (79)$$

6.2 Proof of Proposition 2.1

proof: We need to prove (17) both in the cases of euclidean and supremum norm. We start with the latter, which is easier. Let us set $\|z\| = \|z\|_\infty$. Due to (13), we can write

$$|X_j(z)| < |\tau^{j-1} X_1(z)| + |\tau^{j-1} X_{N+1}(z)| .$$

By expanding X_1 and X_{N+1} as polynomials, we have

$$|X_1(z)| \leq \sum_k |X_{1,k}| |z^k| \leq \|z\|^r \|X_1\|_1 , \quad |X_{N+1}(z)| \leq \|z\|^r \|X_{N+1}\|_1 .$$

Since, for all j , one has $|\tau^{j-1}(z_1^{k_1} \dots z_{2N}^{k_{2N}})| \leq \|z\|^r$, it follows

$$|X_j(z)| < \|X\|_1^\oplus \|z\|^r ,$$

which gives (17).

Let us now set $\|z\| = \|z\|_2$. We are interested in

$$\|(X(z), Y(z))\|^2 = \sum_{l=1}^N (|X_l(z)|^2 + |Y_l(z)|^2) = \sum_{l=1}^N (|\tau^{l-1} X_1(z)|^2 + |\tau^{l-1} Y_1(z)|^2) .$$

Since $X_1(z)$ is a polynomial of degree r one has $X_1(z) = \sum_{|j|=r} X_{1,j} z^j$, where $j = (j_1, \dots, j_n)$ is a $n = 2N$ multi-index. We write

$$\sum_{l=1}^N |X_l(z)|^2 = \sum_{l=1}^N \left| \sum_{|j|=r} X_{1,j} (\tau^{l-1} z)^j \right|^2 = \sum_{l=1}^N \sum_{|j|=r} X_{1,j}^2 (\tau^{l-1} z)^{2j} + \quad (80)$$

$$+ \sum_{l=1}^N \sum_{\substack{j \neq h \\ |j|=|h|=r}} X_{1,j} X_{1,h} (\tau^{l-1} z)^{j+h} . \quad (81)$$

Let us split the estimates of (80) and (81). In the former term we invert the order of the sums over l and j ; recalling that τ acts separately on x and y , we have

$$\begin{aligned} \sum_{l=1}^N (\tau^{l-1} z)^{2j} &= z_1^{2j_1} \dots z_n^{2j_n} + (z_2^{2j_1} \dots z_1^{2j_n}) (z_{N+2}^{2j_{N+1}} \dots z_{N+1}^{2j_{2N}}) + \dots + \\ &+ (z_N^{2j_1} \dots z_{N-1}^{2j_N}) (z_{2N}^{2j_{N+1}} \dots z_{2N-1}^{2j_{2N}}) = \sum_{l=1}^N (\tau^{l-1} z^2)^j < \left(\sum_{l=1}^n z_l^2 \right)^{|j|} = \|z\|^{2r} , \end{aligned}$$

with a rough (but uniform in j) estimate; thus we have

$$\sum_{|j|=r} |X_{1,j}|^2 \left(\sum_{l=1}^N |\tau^{l-1} z|^{2j} \right) \leq \|z\|^{2r} \sum_{|j|=r} |X_{1,j}|^2.$$

Let us now come the second term, (81), and rewrite it as

$$\sum_{|j|=|h|=r, j \neq h} X_{1,j} X_{1,h} \sum_{l=1}^N (\tau^{l-1} z^k), \quad j+h=k, \quad |k|=2r.$$

The idea is to provide again an estimate like $\sum_{l=1}^N |\tau^{l-1} z^k| \leq \|z\|^{2r}$; the tricky point is the possible presence of odd exponent in the multiindex $k = (k_1, \dots, k_n)$. We then first decompose k in its “even” and “odd” parts

$$k = 2k^\natural + k^\sharp, \quad |k^\natural| \leq r, \quad |k^\sharp| = 2(r - |k^\natural|) =: 2s, \quad k_j^\sharp \in \{0, 1\},$$

and consequently decompose the monomial z^k as $z^k = z^{2k^\natural} z^{k^\sharp}$; then we rewrite explicitly

$$z^{k^\sharp} = z_{i_1} \cdots z_{i_{2s}}, \quad i_1, \dots, i_{2s} \in \mathcal{J},$$

where \mathcal{J} represents the subset of those indexes $i_l \in \{1, \dots, n\}$ such that $k_{i_l}^\sharp = 1$. So one has

$$\begin{aligned} |z_{i_1} \cdots z_{i_{2s}}| &\leq \frac{1}{2} (z_{i_1}^2 + z_{i_2}^2) \cdots \frac{1}{2} (z_{i_{2s-1}}^2 + z_{i_{2s}}^2) = \frac{1}{2^s} \prod_{m=1}^s (z_{j_{2m-1}}^2 + z_{j_{2m}}^2) = \\ &= \frac{1}{2^s} \sum_{\substack{l_m \in \{2m-1, 2m\} \\ m=1, \dots, s}} z_{l_1}^2 \cdots z_{l_s}^2 \leq \|z\|^{2s}, \end{aligned}$$

where last sum has exactly 2^s elements. The above upper bound holds also for any translated monomial $|\tau^{l-1} z^{k^\sharp}|$.

From the above considerations we get

$$\sum_{l=1}^N |\tau^{l-1} z^k| = \sum_{l=1}^N |\tau^{l-1} z^{2k^\natural}| |\tau^{l-1} z^{k^\sharp}| \leq \|z\|^{2s} \sum_{l=1}^N |\tau^{l-1} z^{2k^\natural}| \leq \|z\|^{2r},$$

thus collecting the diagonal and off-diagonal elements of $\left(\|X_1\|_1^\oplus \right)^2$ we obtain

$$\|X(z)\|^2 \leq \|z\|^{2r} \left(\sum_{|j|=r} X_{1,j}^2 + \sum_{|j|=|h|=r, j \neq h} |X_{1,j}| |X_{1,h}| \right) = \|z\|^{2r} \left(\sum_{|j|=r} |X_{1,j}| \right)^2.$$

Using a similar estimate for the component Y , we finally get

$$\|F(z)\| \leq \|F\|_1^\oplus \|z\|^r, \quad (82)$$

and the thesis follows. \square

6.3 Proof of Lemma 2.4

proof: We will use (14). By the hypothesis on the seed f we have $f = \sum_{m \geq 0} f^{(m)}$ with $\|f^{(m)}\|_1 \leq C_f e^{-\sigma m}$. Using both the fact that $f^{(m)}$ depends only on the subsets x_0, \dots, x_m and y_0, \dots, y_m , and it decays with m , one has

$$\begin{aligned} \sum_{l=0}^{N-1} \left\| \frac{\partial f}{\partial x_l} \right\|_R &\leq \sum_{m \geq 0} \sum_{l=0}^{N-1} \left\| \frac{\partial f^{(m)}}{\partial x_l} \right\|_R = \sum_{m \geq 0} \sum_{l=0}^m \left\| \frac{\partial f^{(m)}}{\partial x_l} \right\|_R \leq \\ &\leq \sum_{m \geq 0} \sum_{l=0}^m r R^{r-1} C_f e^{-\sigma m} \leq r R^{r-1} \frac{2C_f}{(1 - e^{-\sigma})^2}, \end{aligned}$$

where we also used

$$\begin{aligned} \sum_{m \geq 0} (m+1) e^{-\sigma m} &= 1 + \sum_{m \geq 1} (m+1) e^{-\sigma m} \leq 1 + \int_1^{+\infty} (x+1) e^{-\sigma(x-1)} dx = \\ &= \frac{\sigma^2 + 2\sigma + 1}{\sigma^2} \leq \frac{2}{(1 - e^{-\sigma})^2}. \end{aligned}$$

The same calculation holds for derivatives with respect to the y variables. \square

6.4 Lie derivative of a vector field

Lemma 6.3 *Let $\mathcal{X} = \chi^\oplus$ a cyclically symmetric polynomial of degree $r+1$ and X_F an Hamiltonian vector field of a cyclically symmetric Hamiltonian $F = f^\oplus$, where f is a polynomial of degree $s+1$. Then it holds true*

$$\|L_{\mathcal{X}} X_F(z)\| \leq s \left\| X_F \right\|_1^\oplus \left\| X_{\mathcal{X}} \right\|_1^\oplus \|z\|^{s+r-1}. \quad (83)$$

proof: As already remarked in (15), we can interpret X_F as an r -linear operator, hence there exists \tilde{X}_F such that $X_F(z) = \tilde{X}_F(z, \dots, z)$; we can thus write the Lie derivative of X_F as

$$L_{\mathcal{X}} X_F(z) = dX_F(z)[X_{\mathcal{X}}(z)] = s \tilde{X}_F(X_{\mathcal{X}}(z), z, \dots, z),$$

which, using (82), gives the thesis. \square

Corollary 6.3 *Under the same hypothesis of Lemma 6.3 it holds true*

$$\|L_{\mathcal{X}}^p X_F(z)\| \leq \prod_{j=0}^{p-1} [s + j(r-1)] \left\| X_F \right\|_1^\oplus \left(\left\| X_{\mathcal{X}} \right\|_1^\oplus \right)^p \|z\|^{s+p(r-1)}. \quad (84)$$

Corollary 6.4 *Under the same hypothesis of Lemma 6.3, if $\|z\| \leq R - \delta$, it holds true*

$$\|L_{\mathcal{X}} X_F(z)\| \leq \frac{\left\| X_F \right\|_R^\oplus}{\delta} \left\| X_{\mathcal{X}} \right\|_1^\oplus \|z\|^r. \quad (85)$$

proof: Splitting the term $\|z\|^{r+s-1}$ of (83) in $\|z\|^{s-1}\|z\|^r$, and using¹⁷

$$s(R - \delta)^{s-1} < \frac{1}{\delta} R^s \quad (86)$$

in (83) we get the thesis. \square

6.5 Proof of Lemma 4.1

Let us write¹⁸

$$x_0 = a_0 q_0 + \sum_{m=1}^{[N/2]} a_m (\delta q)_m, \quad (\delta q)_m := (q_m + q_{N-m}), \quad a_m := A_{1,m}^{-1/4},$$

so that $S((\delta q)_m) \subset [0, \dots, m] \cup [N-m, \dots, N]$. Then set $\mathcal{A} := a_0 q_0$ and $\mathcal{B} := x_0 - \mathcal{A}$. Thus we expand the seed x_0^4 as $(\mathcal{A} + \mathcal{B})^4 = \mathcal{A}^4 + 4\mathcal{A}^3\mathcal{B} + 6\mathcal{A}^2\mathcal{B}^2 + 4\mathcal{A}\mathcal{B}^3 + \mathcal{B}^4$, and we deal separately with the five terms of the expansion of the seed. We list below which monomials extracted from x_0^4 , are going to compose the element $h_1^{(m)}$. We assume N odd, the even case follows almost identically. When $m = 0$ we plainly have $h_1^{(0)} = \mathcal{A}^4 = a_0^4 q_0^4$. For $m = 1, \dots, [N/2]$ we have:

$\mathcal{A}^3\mathcal{B}$ term: since

$$\mathcal{A}^3\mathcal{B} = \sum_{m=1}^{[N/2]} a_0^3 a_m q_0^3 (\delta q)_m,$$

with

$$S(q_0^3 (\delta q)_m) \subset [0, \dots, m] \cup [N-m, \dots, N],$$

we take only $a_0^3 a_m q_0^3 (\delta q)_m$. It gives $|a_0^3 a_m| = \mathcal{O}(e^{-\sigma_0 m})$;

$\mathcal{A}^2\mathcal{B}^2$ term: we take

$$a_0^2 q_0^2 a_m (\delta q)_m \left[a_m (\delta q)_m + 2 \sum_{1 \leq i < m} a_i (\delta q)_i \right];$$

by using the additional decay of $|a_i|$, it gives $a_0^2 a_m^2 + 2 \sum_{1 \leq i < m} |a_0^2 a_i a_m| = \mathcal{O}(e^{-\sigma_0 m})$;

$\mathcal{A}\mathcal{B}^3$ term: we take

$$\begin{aligned} & a_0 a_m^3 q_0 (\delta q)_m^3 + 3a_0 q_0 \left[\sum_{i < m} a_i a_m^2 (\delta q)_i (\delta q)_m^2 + \sum_{i < m} a_i^2 a_m (\delta q)_i^2 (\delta q)_m \right] + \\ & + 6a_0 q_0 \left[\sum_{i < j < m} a_i a_j a_m (\delta q)_i (\delta q)_j (\delta q)_m \right]; \end{aligned}$$

by using the additional decay of $|a_i|$ and $|a_j|$ it gives $|a_0 a_m^3| + 3 \sum_{i < l} |a_0 a_i a_m^2| + 3 \sum_{i < l} |a_0 a_i^2 a_m| + 6 \sum_{i < j < l} |a_0 a_i a_j a_m| = \mathcal{O}(e^{-\sigma_0 m})$;

¹⁷defining $x = \delta/R < 1$, we may rewrite (86) as $g_r(x) := rx(1-x)^{r-1} < 1$, which is true since for $x \in [0, 1]$ we have $g_r(x) \leq g_r(1/r) = (1 - \frac{1}{r})^{r-1} < 1$.

¹⁸If $2/N$ one has to set $(\delta q)_{N/2} = q_{N/2}$.

\mathcal{A}^4 term: we take

$$\begin{aligned} & a_m^4 (\delta q)_m^4 + 6 \sum_{i < l} a_i^2 a_m^2 (\delta q)_i^2 (\delta q)_m^2 + \\ & + 12 \sum_{i, j < l} a_i a_j^2 a_m (\delta q)_i (\delta q)_j^2 (\delta q)_m + 12 \sum_{i, j < l} a_i a_j a_m^2 (\delta q)_i (\delta q)_j (\delta q)_m^2 + \\ & + 24 \sum_{i < j < h < l} a_0 a_i a_h a_m (\delta q)_0 (\delta q)_i (\delta q)_h (\delta q)_m, \end{aligned}$$

giving a contribute $\mathcal{O}(e^{-\sigma_0 m})$.

This concludes the proof.

Acknowledgement. We warmly thank Dario Bambusi for many useful discussions and comments. We are indebted with both the referees for their constructive criticism, which helped us to clarify many aspect of the manuscript during the revision. T.P. would like also to thank V.Koukouloyannis, P.G.Kevrekidis and D.Pelinovsky for many interesting and helpful discussions about the present paper, during the AIMS14 Conference in Madrid. This research is partially supported by MIUR-PRIN program under project 2010 JJ4KPA (“Teorie geometriche e analitiche dei sistemi Hamiltoniani in dimensioni finite e infinite”).

References

- [1] D. BAMBUSI, A. CARATI, AND T. PENATI, *Boundary effects on the dynamics of chains of coupled oscillators*, Nonlinearity, 22 (2009), pp. 923–946.
- [2] D. BAMBUSI AND N. N. NEKHOROSHEV, *A property of exponential stability in nonlinear wave equations near the fundamental linear mode*, Phys. D, 122 (1998), pp. 73–104.
- [3] A. CARATI, *An averaging theorem for Hamiltonian dynamical systems in the thermodynamic limit*, J. Stat. Phys., 128 (2007), pp. 1057–1077.
- [4] A. CARATI AND A. M. MAIOCCHI, *Exponentially long stability times for a nonlinear lattice in the thermodynamic limit*, Communications in Mathematical Physics, 314 (2012), pp. 129–161. 10.1007/s00220-012-1522-z.
- [5] R. CARRETERO-GONZÁLEZ, J. D. TALLEY, C. CHONG, AND B. A. MALOMED, *Multi-stable solitons in the cubic-quintic discrete nonlinear Schrödinger equation*, Phys. D, 216 (2006), pp. 77–89.
- [6] C. CHONG, R. CARRETERO-GONZÁLEZ, B. A. MALOMED, AND P. G. KEVREKIDIS, *Variational approximations in discrete nonlinear Schrödinger equations with next-nearest-neighbor couplings*, Phys. D, 240 (2011), pp. 1205–1212.
- [7] C. CHONG AND D. E. PELINOVSKY, *Variational approximations of bifurcations of asymmetric solitons in cubic-quintic nonlinear Schrödinger lattices*, Discrete Contin. Dyn. Syst. Ser. S, 4 (2011), pp. 1019–1031.

- [8] C. CHONG, D. E. PELINOVSKY, AND G. SCHNEIDER, *On the validity of the variational approximation in discrete nonlinear Schrödinger equations*, Phys. D, 241 (2012), pp. 115–124.
- [9] C. CLAUDE, Y. S. KIVSHAR, O. KLUTH, AND K. H. SPATSCHEK, *Moving localized modes in nonlinear lattices*, in Future directions of nonlinear dynamics in physical and biological systems (Lyngby, 1992), vol. 312 of NATO Adv. Sci. Inst. Ser. B Phys., Plenum, New York, 1993, pp. 223–226.
- [10] I. DAUMONT, T. DAUXOIS, AND M. PEYRARD, *Modulational instability: first step towards energy localization in nonlinear lattices*, Nonlinearity, 10 (1997), pp. 617–630.
- [11] P. J. DAVIS, *Circulant matrices*, John Wiley & Sons, New York-Chichester-Brisbane, 1979. A Wiley-Interscience Publication, Pure and Applied Mathematics.
- [12] A. GIORGILLI, *Exponential stability of Hamiltonian systems*, in Dynamical systems. Part I, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003, pp. 87–198.
- [13] A. GIORGILLI, S. PALEARI, AND T. PENATI, *Extensive adiabatic invariants for nonlinear chains*, J. Statist. Phys., 148 (2012), pp. 1106–1134.
- [14] ———, *An extensive adiabatic invariant for the Klein-Gordon model in the thermodynamic limit*, Ann. Henri Poincaré, online first (2014).
- [15] V. KOUKOULOYANNIS AND S. ICHTIAROGLOU, *Existence of multibreathers in chains of coupled one-dimensional Hamiltonian oscillators*, Phys. Rev. E (3), 66 (2002), pp. 066602, 8.
- [16] V. KOUKOULOYANNIS AND P. G. KEVREKIDIS, *On the stability of multibreathers in Klein-Gordon chains*, Nonlinearity, 22 (2009), pp. 2269–2285.
- [17] V. KOUKOULOYANNIS, P. G. KEVREKIDIS, J. CUEVAS, AND V. ROTHOS, *Multibreathers in Klein-Gordon chains with interactions beyond nearest neighbors*, Phys. D, 242 (2013), pp. 16 – 29.
- [18] A. M. MAIOCCHI, D. BAMBUSI, AND A. CARATI, *An averaging theorem for FPU in the thermodynamic limit*, J. Statist. Phys., 155 (2014), pp. 300–322.
- [19] S. PALEARI AND T. PENATI, *Long time stability of small amplitude Breathers in a mixed FPU-KG model*, (2014). <http://arxiv.org/abs/1405.7841>.
- [20] D. PELINOVSKY AND A. SAKOVICH, *Multi-site breathers in Klein-Gordon lattices: stability, resonances and bifurcations*, Nonlinearity, 25 (2012), p. 3423.
- [21] D. E. PELINOVSKY, P. G. KEVREKIDIS, AND D. J. FRANTZESKAKIS, *Stability of discrete solitons in nonlinear Schrödinger lattices*, Phys. D, 212 (2005), pp. 1–19.
- [22] T. PENATI AND S. PALEARI, *Breathers and Q-breathers: two sides of the same coin*, SIAM J. Appl. Dyn. Syst., 11 (2012), pp. 1–30.
- [23] Z. RAPTI, *Multibreather stability in discrete Klein-Gordon equations: beyond nearest neighbor interactions*, Phys. Lett. A, 377 (2013), pp. 1543–1553.

- [24] B. RINK, *Symmetry and resonance in periodic FPU chains*, Comm. Math. Phys., 218 (2001), pp. 665–685.
- [25] ———, *Symmetric invariant manifolds in the Fermi-Pasta-Ulam lattice*, Phys. D, 175 (2003), pp. 31–42.
- [26] K. YOSHIMURA, *Stability of discrete breathers in nonlinear Klein-Gordon type lattices with pure anharmonic couplings*, J. Math. Phys., 53 (2012), pp. 102701, 20.